

1 GCCATGGTGG GGCAGAGGTT GGGAAAGATGG CGTGGCGAGG CTGGGGCGCAG  
51 AGAGGGCTGGG GCTGCGGCCA GGCCTGGGGT GCGTCGGTGG GCGGCCGAG  
101 CTGCGAGGAG CTCACTGCGG TCCTAACCCC GCGCAGCTC CTCGGACGCA  
151 GGTTTAACCTT CTTTATTCAA CAAAAATGCG GATTCAAGAAA AGCACCCAGG  
201 AAGGTTGAAC CTCGAAGATC AGACCCAGGG ACAAGTGGTG AAGCATAACAA  
251 GAGAAGTGCT TTGATTCTCTC CTGTGGAAGA AACAGTCCTT TATCCTTCTC  
301 CCTATCCTAT AAGGAGTCTC ATAAAACCTT TATTTTTTAC TGTTGGGTTT  
351 ACAGGCTGTG CATTGGATC AGCTGCTATT TGGCAATATG AATCACTGAA  
401 ATCCAGGGTC CAGAGTTATT TTGATGGTAT AAAAGCTGAT TGGTTGGATA  
451 GCATAAGACC ACAAAAAGAA GGAGACTTCA GAAAGGAGAT TAACAAGTGG  
501 TGGAAATAACC TAAGTGTGAG CCAGCGGACT GTGACAGGTA TTATAGCTGC  
551 AAATGTCCCT GTATTCTGTT TATGGAGAGT ACCTTCTCTG CAGCGGACAA  
601 TGATCAGATA TTTCACATCG AATCCAGCCT CAAGTGTAT TTCCAATTTT  
651 GTCAGTTACG TGTTGAAAGT TGCCACAGGA AGATATGGAC CATCACTTGG  
701 TGCATCTGGT GCCATCATGA CAGTCCTCGC AGCTGTCTGC ACTAAGATCC  
751 CAGAAGGGAG GCTTGCCTT ATTTCCTTC CGATGTTCAC GTTCACAGCA  
801 GGGAAATGCCCG TGAAAGCCAT TATGCCATG GATACAGCAG GAATGATCCT  
851 GGGATGGAAA TTTTTGATC ATGCCGCACA TCTTGGGGGA GCTCTTTTG  
901 GAATATGGTA TGTTACTTAC GGTCTGAAAC TGATTTGGAA GAACAGGGAG  
951 CCGCTAGTGA AAATCTGGCA TGAAATAAGG ACTAATGGCC CCAAAAAAAGG  
1001 AGGTGGCTCT AAGTAAAATC GGGATTGGAC AGTAGTGGTG CATCTGGTCC  
1051 TTGCCGCTG AGAGCCCCAG GAGACATCGG CTAGAGTGAC CATGGCTATG  
1101 CTCCCGTCTG GAAGATGCCA GCATCTGGCC TCCCACTGTT TTCAGCTGTG  
1151 TCCCCCAGTC CGTGTCTTT TAGAATGTGA ATGATGATAA AGTTGTGAAA  
1201 TAAAGGTTTC TATCTAGTTT GTAAAAAAAAA AAAAAAAAAA AAAAAAAA (SEQ ID NO:1)

FEATURES:

5'UTR: 1 - 26  
Start Codon: 27  
Stop Codon: 1014  
3'UTR: 1017

Homologous proteins:

gi 11066250 gb AAG28519.1 AF197937_1 (AF197937) presenilins int...	668	0.0
gi 8924134 ref NP_061092.1  hypothetical protein PRO2207 [Homo ...	264	1e-69
gi 7303544 gb AAF58598.1  (AE003824) CG8972 gene product [Droso...	186	4e-46
gi 3219925 sp 014364 YB4J_SCHPO HYPOTHETICAL 33.6 KD PROTEIN C3...	69	1e-10
gi 6321538 ref NP_011615.1  Ygr101wp [Saccharomyces cerevisiae]...	64	3e-09

FIGURE 1A

EST:

gi 10216540 /dataset=dbest /taxon=96...	1203	0.0
gi 10215044 /dataset=dbest /taxon=96...	1203	0.0
gi 10212049 /dataset=dbest /taxon=96...	1172	0.0
gi 10154606 /dataset=dbest /taxon=96...	1160	0.0
gi 9141009 /dataset=dbest /taxon=9606...	1144	0.0
gi 9338606 /dataset=dbest /taxon=960...	1094	0.0
gi 9720819 /dataset=dbest /taxon=960...	1090	0.0
gi 5857747 /dataset=dbest /taxon=9606 ...	1033	0.0
gi 10813749 /dataset=dbest /taxon=960...	1009	0.0

EXPRESSION INFORMATION FOR MODULATORY USE:

gi|10216540 Lung  
gi|10215044 Lung small cell carcinoma  
gi|10212049 Lung small cell carcinoma  
gi|10154606 Ovary adenocarcinoma  
gi|9141009 Lung  
gi|9338606 Uterus endometrium  
gi|9720819 Lymph Burkitt lymphoma  
gi|5857747 Colon  
gi|10813749 Dendritic cells

Tissue Expression:

Human Leukocytes

FIGURE 1B

1 MAWRGWAQRG WGCGQAWGAS VGGRSCEELT AVLTPPQLLG RRFNFFIQQK  
51 CGFRKAPRKV EPRRSDPGTS GEAYKRSALI PPVEETVFYP SPYPIRSLIK  
101 PLFFTGVGFTG CAFGSAAIWQ YESLKSRVQS YFDGIKADWL DSIRPKEGD  
151 FRKEINKWN NLSDGQRTVT GIIAANVLVF CLWRVPSLQR TMIRYFTSNP  
201 ASSVISNFVS YVGKVATGRY GPSLGASGAI MTVLAACVTK IPEGRLAIIF  
251 LPMFTFTAGN ALKAIIAMDT AGMILGWKFF DHA AHLGGAL FGIWVVTYGH  
301 ELIWNREPL VKIWHEIRTN GPKKGGGSK (SEQ ID NO:2)

FEATURES:

Functional domains and key regions:

Prosite results:

[1] PDOC00001 PS00001 ASN\_GLYCOSYLATION

N-glycosylation site

161-164 NLSD

---

[2] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE

Protein kinase C phosphorylation site

Number of matches: 3

1	123-125	SLK
2	142-144	SIR
3	217-219	TGR

---

[3] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE

Casein kinase II phosphorylation site

Number of matches: 3

1	25-28	SCEE
2	69-72	TSGE
3	130-133	SYFD

---

[4] PDOC00008 PS00008 MYRISTYL

N-myristoylation site

Number of matches: 10

1	12-17	GCGQAW
2	14-19	GQAWGA
3	18-23	GASVGG
4	22-27	GGRSCE
5	110-115	GCAFGS
6	171-176	GIIAAN
7	225-230	GASGAI
8	228-233	GAIMTV

FIGURE 2A

9 272-277 GMILGW  
10 288-293 GALFGI

[5] PDO00009 PS00009 AMIDATION

Amidation site

39-42 LGRR

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	107	127	1.825	Certain
2	173	193	1.069	Certain
3	226	246	1.654	Certain
4	250	270	1.382	Certain
5	288	308	1.123	Certain

BLAST Alignment to Top Hit:

Alignment to top blast hit:

>gi|11066250|gb|AAG28519.1|AF197937\_1 (AF197937) presenilins  
interacting rhomboid-like protease [Homo sapiens]  
Length = 379

Score = 668 bits (1706), Expect = 0.0

Identities = 327/379 (86%), Positives = 328/379 (86%), Gaps = 50/379 (13%)

Frame = +3

Query: 27 MAWRGWAQRGWGCGQAWGASVGGRSCEELTAVLTPPQLLGRRFNFFIQQQKCGFRKAPRKV 206  
MAWRGWAQRGWGCGQAWGASVGGRSCEELTAVLTPPQLLGRRFNFFIQQQKCGFRKAPRKV

Sbjct: 1 MAWRGWAQRGWGCGQAWGASVGGRSCEELTAVLTPPQLLGRRFNFFIQQQKCGFRKAPRKV 60

Query: 207 EPRRSDPGTSGEAYKRSALIIPVEETVFYPSPYPIRSLIKPLFFTGVFTGCAFGSAAIWQ 386  
EPRRSDPGTSGEAYKRSALIIPVEETVFYPSPYPIRSLIKPLFFTGVFTGCAFGSAAIWQ

Sbjct: 61 EPRRSDPGTSGEAYKRSALIIPVEETVFYPSPYPIRSLIKPLFFTGVFTGCAFGSAAIWQ 120

Query: 387 YESLKSrvQSYFDGIKADWLDsIRPQKEGDFRKEINKWNNLSDGQRTVTGIIANVLV 566  
YESLKSrvQSYFDGIKADWLDsIRPQKEGDFRKEINKWNNLSDGQRTVTGIIANVLV

Sbjct: 121 YESLKSrvQSYFDGIKADWLDsIRPQKEGDFRKEINKWNNLSDGQRTVTGIIANVLV 180

Query: 567 CLWRVPSLQRTMIRYFTSNPAS----- 632  
CLWRVPSLQRTMIRYFTSNPAS

Sbjct: 181 CLWRVPSLQRTMIRYFTSNPASKVLCSPMILLSTFSHLFHMAANMYVLWSFSSSIVNIL 240

Query: 633 -----SVISNFVSYVGKVATGRYGPSLGASGAIMTVLAAVCTKIPERLAIIF 776  
VISNFVSY+GKVATGRYGPSLGASGAIMTVLAAVCTKIPERLAIIF

Sbjct: 241 GQEQFMAYVLSAGVISNFVSYLGKVATGRYGPSLGASGAIMTVLAAVCTKIPERLAIIF 300

**FIGURE 2B**

Docket No.: CL001058DIV  
Serial No.: (to be assigned)  
Inventors: Jane YE et al.  
Title: ISOLATED HUMAN PROTEASE PROTEINS, ...

Query: 777 LPMFTFTAGNALKAIIAMDTAGMILGWKFFDHAHLGGALFGIWYVTYGHeliWKNREPL 956  
LPMFTFTAGNALKAIIAMDTAGMILGWKFFDHAHLGGALFGIWYVTYGHeliWKNREPL

Sbjct: 301 LPMFTFTAGNALKAIIAMDTAGMILGWKFFDHAHLGGALFGIWYVTYGHeliWKNREPL 360

Query: 957 VKIWHEIRTNGPKGGGSK 1013

VKIWHEIRTNGPKGGGSK

Sbjct: 361 VKIWHEIRTNGPKGGGSK 379 (SEQ ID NO:4)

Hmmер search results (Pfam):

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
PF01694	Rhomboid family	23.3	1.8e-05	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF01694	1/1	201	292 ..	59	147 ..	23.3	1.8e-05

FIGURE 2C

1 CGAGGTTTCT TCATGTTGGT CAGGCTGGTC TCGAACTCCC GACCTCAGGT  
51 GATCCGTCCG CCTCAGCCTC CCAAAGTACT GCTGGGATTA CAGACGTGAG  
101 CCACCGCACC CGGCCTTAT CTTTCATTTT TTTTCATGTA TTTTCCTTTA  
151 TTTTAATCAC TTATCCAGA AACATATCCT CGTCTTGACA GTGCTGTGGT  
201 GCCTGTGGTT TCCAGAAGCT GGGTGTGCTG TGTGTCTGTG GTTTGAGGAA  
251 GTTGGCCATG GAACTGACAG AGGAAGCAGA GTAGTCGTTG CCATTTTCA  
301 GCCTAGTAGG CAGGATCAGG GACCCATCT TGCTCTCTT GCCTTGAACC  
351 ACAATTAGAA TAAAACACCA AAGCCCTGAC TGATCATGAT CATAGCAATC  
401 CGATCTTAT GATCATGGCC AGACCATTCT CAGGTCGTCT TTACCCCTAAG  
451 ATATCAATCA CTGGGTATGA CAACCTAGAC CTAAGGGTGC ACTCTGGGT  
501 GTAAAGATGA TTAACTCTCC CAAAGGAATC TAAGGAATCC AGAGCAACAC  
551 GAATCACTGC TCTCTTCTA TAGGGTAAAC CTCCCAAGAC TCCAGTCCCT  
601 GTGAGGAGGC TCTGCCCGCC TGCCCTTCCC AGGGTTCAG GCTCCACATT  
651 GGGAGGTGTA CACAGTGCCTC TTGCTCTTC ATTGCCTTGT GTATGATCCC  
701 TTTTCCCATC TTTGCATAAA TGCTGTCCCT CTCACCACATCT TTAAAAGAGT  
751 TCTGGGTAAT TATTTACAA AGGTGGTATA ATGCTGTAC AGTCCCTGCT  
801 AGTGAGACAT CTGATACAAAC TGATGGAATC AGTTCAACAA AATGCAGTAA  
851 AATTTTATTT AATGTACTAC GGAGAAAGAA AAAATGCTAC CAGTTATAAG  
901 ATGCATCCTG ATTTCAGATA TTAAAATGGA AAAATGTCT TAAGATCTGT  
951 GAAAAATGTA GCTTCCTTTC CCACCTCTCA AGTGGGAGAG CAAAAACTGG  
1001 ACAGACTAGA AATGCCAGGG GCTAGCTGAG AACCTTACAG AATGAGCAAC  
1051 TGCAGGAAGCC ACAGGTAAAC CCGAGATGTA GATCAGCTGC CAGGGACAAG  
1101 ACAAAAGAATG TTTTCTAAAG TAAATCCTCT TACCAAGTATG TTATTGAAAT  
1151 CAGTCCTTAT TGGCATCGAA GAAGGTGAAA GTGCTACTTG CCTGTTGCCT  
1201 ACAGAGACTG GAGGAATGAC AAATGTTAA ATTATTTAA TTCAACAAAGT  
1251 AGAGGAATAC CTGCTATGTG AAGGAGTTGT GGCAATTCTA AAAATTAAATA  
1301 TATTTTGTGAG AGTTGTAGT TTCAATAAT AATTTCTTAT CTAAAATGTA  
1351 ACAAGTTAAT TATATTATCG AATAAACCTC ATTTCTGTAG TACTAACAAAC  
1401 ATCACACCTT ACAGAAAAAG GAAAGTCACT CAACTCCCAC ATGTAACAG  
1451 ACTTTAGAAG CAGTTGCAGA GGTTTCTAA ATTATCCTG AATTCCCTATC  
1501 ACATGACTAT TTTTCTCAGA CATGTTGACC TTCACCTACA CAGATGACTC  
1551 ACATATGTTT CCATAAGCTG GCAGTAAGTT TAAGAAGCAT ACCATGCCCT  
1601 GAGGAAAAAG AAGTAATGTT AGCTCTTCTA CTCTGGCCA AAGAACCTAA  
1651 TTCTGTATAT TACTTCTGTC TTTGGTTTGG CTATTATAGA CAATAAATTA  
1701 TTGATCTGAT TATAATTGAG AAAAGTAAGC TCTTCTAAAG AAGTAAAATA  
1751 TGGATCTAGG GAAAGGAAGT TAGCTCCAG AGCATTACA ATTTCCAGG  
1801 AATTCTGTGA CTTTACCAAC CCTAGGCAGT GCTGATACTT TAAAAGCATT  
1851 CATTCACTT GCTTTTTTT GGCTCACCCCC CTATCCCCA GGTATAACAGT  
1901 ACTCTTACAT AATTGTGGAA GAATCTTACA AGGGGGTAAT GTAGATCAGA  
1951 CTTTCCGTCT TTCACTTTA ACCTCCCTAA ATTATAAATA TTTATTTGT  
2001 AGGTATTATA GCTGCAAATG TCCTTGTATT CTGTTTATGG AGAGTACCTT  
2051 CTCTGCAGCG GACAATGATC AGATATTCTA CATCGAATCC AGCCTCAAGT  
2101 AAGTCTAACT TGTGTGAATT TATTTAAGG TAGAAATAAT ATGAAAGAAA  
2151 TATGCTTGTAG TTAATGGAAG TGCTGTAAAA AAGACGAATT ACCTATCAAT  
2201 AGCTACAAGC AAAATGCAGA GGATAGGCTG TAAGCTCCTT CACTGAGGAC  
2251 AGGGACCTCA CCTCTCTTT TCTTTTCTT TGTTTTTTT GAGACGGAGT

## FIGURE 3A

2301 CTTCTCTGT TGCCCAGGCT GGAGTGCAGT GGTGCAGTCT TAGCTCACTA  
2351 CAACCTCAC CTCCCAGGTT CAAGTGATT TCCTGCCTCA GCCTCCCTAG  
2401 TAGCTAGGAT TACAGGTGCC CGCCACCACA CCCAGCTAGT TTTGTATTT  
2451 TTAATAGAGA CAGGGTTCA CCGTGTGGA TAGGCTGTC TTGAACACCT  
2501 GACCTCAGGT GATCTGCCTG GCTCGGCTGG AGTGCAGTGG CGTGATCTCA  
2551 GCTCACTGCA AGCTCCGCCT CCCGGTTCA TGCCATTCTC CTGCCCTCAGC  
2601 CTCCTGAGTA GCTGGGACTA CAGGTGCCCG CCACACGCC CCGCTAATT  
2651 TTTGTATTT TTAGTAGAGA CGGGGTTCA ACATGTTAGC CAGGATGGTC  
2701 TCGATCTCT GACCTCGTGA TCCGCCCCGC TCAGCCTCCC AAAGTGCTGG  
2751 GATTATAGGC GTGAGCCACT GCGCCCGGCC AATTACTTT TTATTTTATT  
2801 TTATTTTATT TTTTGAGACA GGGTCTTGCT CTGTTGCCCA GGCTAGAGTG  
2851 CAGTGATACG ATCTTGCTC ACTGCAACCT CTGCTTCTCA GGCTCAACTG  
2901 ATCCTCCAC CTCAGCCCC AGGAGCTGGG ACTACAGGTG CATGCCACCA  
2951 TGCCCAGCTA ATTTTTTTG TTTTAGTGC AGATGAGGTC TTGCCATGTT  
3001 GCCCAGACTG CTTATTTTT TCTAATCAAC TTTGCCATA AGGACAAGTT  
3051 GCTTCATTG AACTGAGAGT TTTTATTGGT TGCTTACTAA GTAGAAAAGA  
3101 ATATTTATTA AGACAGCTT TTGTCACTTT TAAAAATGAT GTCTTAAGCT  
3151 GGGCATAGTG ACTCACATCT ATAATCCCAG CACTGGGA GGCTGAGGCA  
3201 GGTGAACTGC TTGAGCTCAG GAGTCGAGA CCAGCCTGGG AAACATGGTG  
3251 AAACCCCATC TCTACTAAAA ATACAAAAAT TAGTTGGCA TGGGGTATGT  
3301 ACCTGTGGTC CCAGCTACTC AGGGAGGCTG AGGTGGGAGG ATCACTTGAG  
3351 CCCTTGAGCC TCAACTTGAG GAAAGTTGAGG CTGCAGTGAG CCAAGATCAG  
3401 TGCCACTGCA CTCCAGCCTG GGGCAGACA GCAAGACTCT CTCCAAAAAA  
3451 AAAAAAAAGT CTTAAAAATA GCTGTTTTG TTTCCATGT TTGTTTCATA  
3501 AATTTTTTTT TTTTGAGATA GAGTCTCGCT CTATGGCCA  
3551 GGCTGGAGTG CAGTGGCTCA ATCTTGCTC ACTGCAAACCT CTACCTCCTG  
3601 GGTCCAAGTG ATTCTCCCGC CTCAGCCTTC CGAGTAGCAG GAATTACAAA  
3651 CGTGCGCCAC CACACCTGGC TAATTTTAT ATTTTAATA GAGATGGGGT  
3701 TTGACTATGT TGGCCAGGCT GGTCTTGAAC TCCTGACTTA GTGATCCGCC  
3751 TGCCCTGGCC TCCCAAAGTG CTGGGATTAC AGGCGTGAGC CACTGCGTCC  
3801 GGCTTAATT TAAAAGTTA AAATGGATAA TTTTATTGG CTGTGTGTTT  
3851 CATGATTACC AGACTATGTT TCTCTCTCTT GTAGAGGTCC TTTGTTCTCC  
3901 AATGTTGCTG TCAACATTCA GTCAATTCTC CTTATTTCAC ATGGCAGCAA  
3951 ATATGTATGT TTTGTGGAGC TTCTCTTCCA GCATAGTGAA CATTCTGGGT  
4001 CAAGAGCAGT TCATGGCAGT GTACCTATCT GCAGGTAATA TGCTTAATC  
4051 TCGGGCCCT TGAGAGTATA AGCACTCTAA GCTATCTGCA GAACGGACAA  
4101 AGGAATGAT TACTGCCATA TTCTACACGT AGTGAGTGCT CAGAACATAT  
4151 TTGTTTCTCA CAGTGTATGT AGAGAAGGGG GCCACAGATT GGTGGAGATG  
4201 TTGCTTTTC TGTTCATTTT GCTGATTCT TCTTACATAT GAATTATGTG  
4251 GGTATGTTA ATTTAAGTT AGGATAAAC GGCCTTAAGT AAGGGTTAGT  
4301 GTAGAATTAA AGCATGTCAT TTTGTAATC TCATGGGCC TTGATTTCAT  
4351 TAGTTAGGC CCTCCATTAA ATAGATAGTG GTTCCCAGAC TTCCCGGCTG  
4401 CCTCAATCTC CTGGGTCTT GTAAATAAC CTTAAGCAAG CTCAATTCCC  
4451 CCAGTGTGTT CAGTCACAG AAAGTTAA ATCAGAGCTA TACAATATGA  
4501 TTGTCAGAG TGAGTTGTT CTGCTTCTT TGCAAGAATG TAGCAGGGAA  
4551 CCACTCCTA GCCATGGTCT TGAAGATGGT ATCGTTCTT ATTCAGTTA

FIGURE 3B

4601 GGAAATTCTC ATGCATGAAT CCAGGTCCT AGATGCTGCT AACGTGACAG  
4651 TTGGTCAAAT TTTACTTACC TCTCTGTTG TAAAATGTAC TTACTTAATA  
4701 CAATATAAAA ATTAATTCTC AAAATCTCTA CATTAGAAA CAGTATATCT  
4751 GGCAGTTGTG CTGTGATGTA GTGAAAAACA CTAAGCTTGG CGATAGACCC  
4801 AGGTTCAGAT CCTATTTCTA CTACCAGCTG AGTGTGTTG CAAAAATGAC  
4851 TAAACCTCAT GATACTTACC TCCTCATGAC AAGGGTTAA AGAAAGGACT  
4901 ACATAAAAGC ATCTACCACA AGCCCCAGAG TAGATGCTTA ATTAGTGTTC  
4951 ATCGAATACT TATGTGTATC TAGTCCTTCA AAAAAAGAAG CTGAGCATTG  
5001 TGTGGCTT GTAAGATAAG TGTATAGTT TTTCCAAGC ACTAGTTATG  
5051 TTGTAGTTAC AGAGGGTCTG TTTCAGATAC ATTAATTCT GCTCCATAGG  
5101 AGGTTTTAA AAATGAGCCA CGTTGACTCA AATGGCACTG AAGCCAAAGA  
5151 GACTTACGGG ATCATCCAGT CTGTTGTCCC ACCCCAGATA TTCTGATTT  
5201 GTGTGTCCTGG AGTACAGCCA GAGAATATAC TCTTGGGAAT GAGTCTTCAT  
5251 GTTATAGTTG AGGAAATGG TAACTGAGAA GTGGAGTGAA TGACCGTGT  
5301 GCTCAGCAGA TCATGCAGCA GGTCAAGACTT TTCATCCCCT GTAAAGTCGC  
5351 TGAAATGATA GGCAGGAGAA GTATTCATGC CCGTACCCCT ACAGTGATCC  
5401 AGATTGAAAC CCGACACTGT TTATCTGTGT AGAAATCAGA AATGAAAACC  
5451 ATTTTCATGG CTGGATGTGG TGCCGCACGC CTGTAATCCC AGCTACTCAG  
5501 GAGGCTGGGG GACAAGAATA ACTTGAACCC GGTAGGCAGA GGTGAGTG  
5551 AGCCAAAATT GTACCACTGC ACTTCAGCAG CCGGGGCGAA AGAGTGAAAC  
5601 TCTGTCTCAA AAAAAAAA AAAGAAAAG AAAAAAAAAG TAAACCATT  
5651 TTATACCTCA CTTAAATTAT TGTAATGTGA CTTGTTTTTC AGGTGTATT  
5701 TCCAATTITG TCAGTTACGT GGGTAAAGTT GCCACAGGAA GATATGGACC  
5751 ATCACTTGGT GCAGTAAGTA TTTCTATTGT AAATTTTTT TAATTTAATT  
5801 TTTAAATTAA CTTTGAATA AGTTAGACT TAGAAGAATG TTGTAAAATT  
5851 GATAAGTAGG TTCTCATATA CCCTTCACCC TACTGTAAAC TAACATCGAA  
5901 ACCAAGAAAAT TAACATTGAA ACAATACAGT TGACTAATT AGAATTATA  
5951 CATTGTAAA GCTTGTAAA TGTCGGCTA TAGCTTTAA CCATTGGTCA  
6001 TATATATATG TTTACCAGAG CAGAGTATAT CTCAGAACAG TAAGTGTGCA  
6051 ATCCTCGTAA ACCAGAGAGC CTAATCCAGT ATTGGAAGAT TCTAATTATA  
6101 GATTGAATC TGGTACTTTA TCCTCTTATT TAGTCATAT TGAGTGCCT  
6151 ACTAGGTGCT ATGCTAGAGC CTGGGGATAA CAGCTGGTGA GCAAGATGAT  
6201 CACGATTATT TGTGTTGGTT TTAGAAAGTG GGGAAACAACA ACAACAAAAA  
6251 AGGCTCCTGC CCTCAGAGCT CTTATATTCT GGATGCTTAA AAAAAATT  
6301 CTTAGGCTGG ATGCAGTGGT TTACACCTGT AATCCCAGCA CTTTGGGAGG  
6351 CCAAGGTGAG AGGATGAGCC CAAGAATTG AAACCAGCCC TGTAACATA  
6401 CCAAGATCCT ATCTGTACAA AAAAAATTAA AAAATTAACT GGGGGTGGTG  
6451 GCTTATGCCG GTAGTCTCAG CTACTCAGGA GGCTGAGGAA GGAGGATAGC  
6501 TTGAGCCTAG GAGGTTGAGG CTGCGGTGAG CTGTGATTGT ACCACTGCAC  
6551 CCCAGCCTGG GTGACATAGC AAGACCCTAT CTAAAAAAA AAATTTTTT  
6601 TTAAGTGTGT TTTGAGGCTG GGTGCAGTGG CTCACACCTG TAATCCCAGC  
6651 ACTTTGGGAG GCTGAGGTGG GCAGCTCACT TGAGGTCAGG AGTTCAAGAC  
6701 CAGCCTGGTC AACATGGTGA AACCTGTCC CTCCGTAAAA TACAATAATT  
6751 AGCCAGGTGT GGTTGTGCAT GCTGTAAATC CCAGCTACTC GGGAGGCTGA  
6801 GGCAGGAGAA TTACTTGAAC CCAGCGGGTA GAGGTTGCAG TGAGCTGAGA  
6851 TTGACCACT GCACTCCAGC CTGGGTGACA GAACAAGACC CTGTCTCACA

## FIGURE 3C

6901 GAACAAGACC CTGTCTCAA GAAAAAAAAT TTTTTAAGT GTCTTTGAG  
6951 TTTAATGGCA GATTTCTGGG CACATGGAAA TCTTTATGTA ATATTTCTT  
7001 ACACATTCA G TTTGTACTTA TTTAAATACT AATTCACTTA AATGCATTCA  
7051 AATAGGGAAT TCCCTATTTA AAGGAACCTCT AAAAAGGTCA ATTTTGAAAAA  
7101 GAATTCTTAT GTAAAATAAC CATTCCCTAA TTTGTATGTT CCCCAAATT  
7151 GTTACACTT AATTTCTTA GTGAGGCCTG TGTTCTGTCC TGTGACCACA  
7201 TGCTTCTTA AGCCTCCCTT TTTCCCTTCG TGGAATGTTT ATTTCTTTA  
7251 TACAATTCG CTCTGATATA ATTTATATAT TTCAATCAT ATTTGTCTACC  
7301 TCATTCAACA GCTAAGCACC TAATATATGA AGGCAGTGAA GACCACTAGG  
7351 ATGAATCAGA GACTCAGAAT TCGAATTTAG CTGGGGAGAA AACATGCACA  
7401 CATCTAATAC AACTGAAAG GAATGAGGAT TCTCTAGAGG ACTTTGGGG  
7451 CTCTAAGAGT GAAGAGACCT TTCTAATTAG CTGAAAGGAC CTGCGAGGGC  
7501 ATTTGATGT GCTCTGGAC AGCTGTTGTC CTCATCTTAT AGATAAGAAA  
7551 CTGAAGTGCA AACTTAATGA AGTATGGCAG TAAGGTATTT GGAGTTAGAG  
7601 TGGGGGTGAA TCCTGGTTCT GCTACTTACG TGTGATTCT AGGACATATT  
7651 ACTGAACCTC TCTGAATTTC AGTTTCCCTT TATAAAATGG GGATAACACC  
7701 ATCTATTTCT GAGGTGCAAA GCAAGTACAT TTAGAGTGCT TAGCACAATA  
7751 AGAACACAT GGTAAGAAAT GTGGACATGG TAGTTCCGTG TCAGTCATCA  
7801 AAATCCTACA GCGCCGTGGT AGGATAACAT TATCCCCAA TATCTTAATG  
7851 AATCTGTGAT TAAAATTCAA GGAAATTAAA TCACCAAGTA TAATGGCATT  
7901 TTTAATGAGA AATCTGGAA AAAAACACCA TTAACAAAGT TGTGTTGTTA  
7951 CAAAATGTAA AGCGTTAGTC CTCTGGTTT AGTGAGACGT TATAAGATGC  
8001 AGGGGACAGC CAGGCACAGT GGCTCACGCC TGAGGGCCA ACACCTTG  
8051 AGCCACGGCA GGAAGATCAC TTGAGCCAG GAGGTTGAG ACTAGCCTGG  
8101 GCAACAAAGT GAGACCCAT CTCTACAAAA AATTTCAAAA TTAAGCCGG  
8151 CATGGTGGCA TGCACCTGTA ATCCTACCTA CTCAGGAGAG GTGGGAGGGT  
8201 GGGAGGAATG CCTGAGCCTA GGAGGGTGAG GCTGCTGTGA GCCATGAGCA  
8251 TGCCACTGTG CTCCAACCTG GACAACATAG CGAGACCCCA TCTCAAAAAA  
8301 AAAAAGAA AGTTGAATGG GACTGTTAAA ATATGTTGT AAATTACTGT  
8351 ATTGGTACTA TCCTGGATAA TTTTTAAACT TTTCTGTAGA GACAGGGTCT  
8401 CCCTATGTG CCAAGGCTGG TCTCAAACCTC CTGGGCTCAA GTGATCCTCC  
8451 TACCTGGGCC TCCCAAAGTG TTGGGATTAC TGGTGTGAGC CACTACACCC  
8501 GGCCAATTGT CTTTCTTAT TCAAGTTGAG ATTTTCTGG TTCTTGATAT  
8551 GATGAGTGTAT TTTTCAGTTG AAGCCTGATC ATTTAGATA TGATGAGACT  
8601 TTGGATCTTA TTGAAATCTG CTGTTTCAGT GGTCTTCCTC TGACACTGTT  
8651 CTGATGAGGA GAGGGGGTGC CGTGAACCTGT TACTGCTGGG TGTAGGAGTA  
8701 GACGTCCAGG TTCTCACTC AGCCGCCTT GCCTCCTGAG TGATAGGGC  
8751 TCTTGTCACT GCAGGGCAGG GATGGGAGCT GAGGGCGTGC AGGCTACCTA  
8801 GTGTGCCTCT GCTAATGTG CTGTTGCTAG GAGGAGCAAG GGTGCTTCTT  
8851 TCCGCTGACA CCGCCTGTTA GCGTATTGG GATGCCTCAT TACAGTGTGG  
8901 CAAGGGTGGG AGTCTAGGCT CTGCTCAGCC TTTGCTGGC ACCCGTTCT  
8951 CTAAATATTG TCTAAAAGGT CTCTTTGCT AGGCTATCTT TTTTTGGTCC  
9001 TTGACTAGAG AGAACATGTT GAGGGATGAT CGATATGAGG CCAAAAGAAA  
9051 GCCCAGGGAA CTCACCACCA CAACATTGAT TGAATCTCA GCTTCCTAGC  
9101 TGGTCCGCTT TCCTCTCTC TTCTTCACTA GTCCTCTTAC ATTTGTTCA  
9151 TATGTAACAC CCAGGGCTT TAGCTGTACT TAGCTTTGT AAGCAGAGGG

## FIGURE 3D

9201 AGCAGATTCA CTTAAATTAT AATACCAAAT AAAGTTAAAA AACATAAGTA  
9251 TGATAGATT TGAGATTATA TAGATACAGA AAAATGTTG TGAGCCCAGG  
9301 CGCAGTGGCT CACAACGTGA ATCCCAGCAC TTTGGGAGGC CGAGGTGGGT  
9351 GGATCACTTG AGGCCAGGAG TTGAAACCA GCCTGGCAA CATGGTGGAA  
9401 CCCCCATCTCT ACTAAAATA CAAAAATTAG CTGGGCATGG TGTTGTGTAC  
9451 CTGTTAGTCC CAGCTACTTG GCAGGCTGAG GTGTGAGAAT TAACTTGAAC  
9501 CTGGGAGGCG GAGGTTGCAG TGAGATCGTG CCACCGCACT CCAGTTGGG  
9551 CAATAGCGAG ACTCTGTCTC AAAAAATATA TGTTATGAA ATAAGTAAAA  
9601 AAAATCAGA TGTGCATT GATTACAGGT ATATAACCAG TACATAAAAA  
9651 TATTGATGGA GAACAAAAGA CCTTCACCTC TTCCCATGGA CCCACACCTC  
9701 TTAGGTCTGT TGGATCAGGG TTCACTGACTC ACTGTACTTA AACTGTGTAT  
9751 GAATGTGAGC GTTTCTGAG AAGAGAAGGG TTCATTTCA TAAATTCTT  
9801 CTTTCTGACT CGAAAAAGTG AAAAAAGTCT CTCTGCATGG GAGTAAGCCC  
9851 AAATATTTGT CAAAAAACAA GTTGTGATT ATTCAAGACAT ATAATATTT  
9901 AAATTTATAT AAAAGCCACA TCGAGAAAAT TCTAGAAGGA TGATGGAACT  
9951 GTGTATGTAA TAATTACAAT AAGTTATAAT CACAAAAAAA CCAGCGTTCC  
10001 ATGGAATTGT ACAGATAACG ACAATTTTT TTAACAGATG GAGAATAATC  
10051 ATCTATGGAA TAGTAGTTA GAAGAACTTC ATAGAATTT TTTTTTTTTT  
10101 TTTTTTTTTT TTTTTGGAG AGGGAGTTTC GTTCTTGTG CCCAGGCTGG  
10151 AGTGCAAAGG TGCGATCTCG GCTCGCTACA ACCTCTGCCT CCCGGGTTCA  
10201 AGCGATTCTC CTGCCTAAC CTCCTGAGTA GCTGGGATTA CAGGCATGCA  
10251 CCACCATGCC CAGCTAATTT TGTATTTTTA GCAGAGACTG GTTCTTCA  
10301 TGTTGGTCAG GCTGGTCTCG AACTCCAGAC CTCAGGTGAT CTGCCCGCCT  
10351 CAGCCTCCCA AAGTCCTGGG ATTACAGGTG TAAGCAGCTG TGCCCTGGCAG  
10401 AACTTCATAG AATTTTAATG CTCTTTATA TCAACTAATC AAATTATATT  
10451 TGCTTCATTT TGGGAAACG TGTAATTTTG ATTTGTTTG GGGTTTTTTT  
10501 GAGATAAAAGT GTCACTCTGT CGCCCAGGCT GGAGTACAGT GGCTCAATCT  
10551 TGGCTCACCA CAACCTCAGC CTTCCGAGTA GCTGGGACTA CAGGCGCCA  
10601 CCACCACGTC TGGCTAATTT TTGTGTTTTT AGTAGAGACG GGGTTTCACT  
10651 ATGTTGGCTA GGCTGGCTT GAACTCCTGA CCTCAGGTGA TCCACCTGCC  
10701 TCGGCCCTC AGAGTGTGG GATTACAGGC GTGAGCCACC GTGCCCGGCT  
10751 ACAATTATAG TCTCTTGAC AGAAGCCAGC TTGGTCAAAA TTCAGGTCTT  
10801 CTTGGGTCCT CCTTTGAGG AGTGTTCATG CTGTCCTTCC ATCTTGAGT  
10851 TACCCCTGACT TCTAAGAATG CAACCCGAGC TTGTTCCCT GTTGAGGCCA  
10901 CTTGGCAGTT ATATGAGGGA CTGGGGACAT CTGAGATCTC TGGGACTCAT  
10951 AATAATTTTC TTAAAGTT TAGTAATTCC CCAAATGAA GATAATCTT  
11001 TATTCTGAAG CAACCCGTCA CATAGAAGAC ATTAAGAAAA CATTGATTAA  
11051 GAGAGGTAGA TGCTATTTTC CAGAAACAAAC CGTTTTATA TGAAAAGGTA  
11101 GGAACCTTTC TTTTAATGA TAGGGGCTTC TTTCAAAGT TATTTGCTC  
11151 TTAGGTGTCT TTTTTTTTTT TTAAACATC TCATTCTAA ATAATTAAAA  
11201 ACTTATGGGA AAGTTGCAGG GAATAGTACA GAGGACTCCC ATAAAGTCTT  
11251 TTTTGTGTT TTGTTTGTT TTGTTTGAG ACAGAGTCTC GCTGTTTAC  
11301 CCAGGCTGGA GTGCAGTGGG ACAATCTCGG CTCACTGCAA CCTCTGCC  
11351 CCGGGTTCAA GCAATTCTCG GGCCTTAGCA TCCTAAGTAG GTGGGATTAT  
11401 AAGCATCCGC CACCACGCC AGCTAATTTT TTTTTTTTTT TTTTTTTTG  
11451 TATTTTGTAGT AGAGACGGGG TTTTACACG TTGGTCAAGC TGTTCTCAA

## FIGURE 3E

11501 CTCCTGACCT CAGGTGATCC ACCTGCCTCG GCCTCCAAAA GTGCTGGGAT  
11551 TATAGGCGAG AGCCACTGCA CCCAGCCCCA TGTAGTCTTT TAAAAAAGCA  
11601 GGCAACTCAG GTTTACTAGT TAACATGCAA AAAACTGCAC ATATTTAAAG  
11651 TTTGGTAAGC TTTGACATGT AGACACCCGT GAAACCATCA CCACACTCAA  
11701 GATCATGGAC ATATTCATCC CAAAAGCTTC CTAGTGGTCA CTCCCTCCTG  
11751 CCCCTCCTCT ACCCCTGGCG ACAACTTACC TACTTCTACT AAAGATAAAAT  
11801 TAGTTTGCAA ATGGAACCAT ACAGCATATA CTAGTATTTG TTGCTCTGGC  
11851 CTCATTTACT CTGTATAATT ACTTTGAGAC TCATCCATGT TCTGTGTATC  
11901 AGTTTATTCC TTTATTATTT TTGAGACAGG GTCTTACTCT GTTGGCCAGG  
11951 CAGGAGTGC A GTGGTGAAT CATA GCTCAC TGTAACCTTG ACCTCCTGGG  
12001 CTTAAGGGAT CCTCATGCCT CACAATGTGC TGGAAATTACA GGC GTGAGCC  
12051 ACCACACTGG CAATGTTTG TTTCTTTATG AAGATGAATA AAGATTTCAC  
12101 ATGAATTTTT TAAGATGAAA CATGCTTCAT GCATGCAGGT TTCTTTGGC  
12151 GTATTCATGC CCACTCCCTC TGGTTGGAGC TTTGTCAAG AAGTGTGAGC  
12201 AGTTCTTTC TAGGCCATAG GTGAAAGATG CGCATGACAC GCTTAGCACT  
12251 GTCCTTGCAGG TT CATGAGGC ACATACATCT TACTGCCCG TAGTAAAAAT  
12301 TCAGTCTTC CAAGCGATTA CTGTGTGAAG GACATTAGT TCCTTCACCT  
12351 ATTATTGGGG ACATAAGTAA CTGAAAGCTT TGAAGCTTTG TGCTCACCTA  
12401 GAAATGTGCA GCATGTAAAC TTTCTAGAAA ATGTGCTGCT CTTAGACCT  
12451 TGTAGCCACT AAGCAGTTGC ATATTGAGTT TCCCATTCTC CCTGCTGTGT  
12501 TACTTTGCAG TCTGGTGCCTA TCATGACAGT CCTCGCAGCT GTCTGCACTA  
12551 AGATCCCAGA AGGGAGGCTT GCCATTATTT TCCCTCCGAT GTTCACGTT  
12601 ACAGCAGGGAA ATGTAAGTAT TTTTATGAAG TGCA GTGCTGCTG GGGATAGTGG  
12651 TGATGTTTTT ATGTTGAGTG GGTTCTTGCC CTTAAGTTAG AAATGTCACT  
12701 GCTGGAGCAA TCACAGTTGT GCGCCTTGTT TCTTGCTGCC TTTCAGGCC  
12751 TGAAAGCCAT TATGCCATG GATACAGCAG GAATGATCCT GGGATGGAAA  
12801 TTTTTTGATC ATGCGGCACA TCTTGGGGGA GCTCTTTTG GAATGTAAGT  
12851 TTGAGTGTAA TTGATTGCTA AACTGCTTCC TTGGGTCTG CGCTCCTCCT  
12901 ACCCCAGCCT CACCCCTACC CCCCCTCCCC ATGGCAGAGA CATTGAAC  
12951 TGCAACGGAA GCAGAAGCAG GTGGGCTTGG GAGGGTGAGG AAACCTCAAC  
13001 ATGGCTTGCT TTGGGTTTAC CCAGCATACC TGGCTCATTG TAGAGACAGT  
13051 CTGTGCCTT ACCCTACGCT TAACCTTAAG TTGCCCCAAC TGTGGCCTG  
13101 TTATTCCCAG CCCCCCTTTA GAAGACTGCA GCCTGGCCCC CAGTCTATGC  
13151 TGACATCTTC TTTTCTCCCT TCAGACTTTC CTGCCCTCCT CTCCCTGCC  
13201 TGGCGTCCCA CCCTGCTACC CTGACCTCTG TCTCGCCAGT GCTATTAGA  
13251 CATGCTGAGT TGGCGGAGCC ATTGCTCTGT ATGACTGGAG TAGAGGCCGG  
13301 TGACTGCAAA CCAATGTGGA CCACCTACTG AGTACCCGCT GTATGCAGGC  
13351 ACCAAGCTAG TTCCCTTATG TTATACTATT ACTACTCCC TTTTACTGAT  
13401 GGGAAACTGA GGCTCAGACA TCATCTTCCC CAGGCCAAC AGCTCTCAA  
13451 TAGCAGAGCA GAGCTGAAA CCCACCTCTA TAAGCCCTT CCACCCCCAC  
13501 CACACCATAT GGAATTGGTT GCTAAACTGC TTCCCTGGGT CACAGCAAAT  
13551 GGCATTGTGG TTACAAGACC TTCCACGTGT GCTTCAAACCA ATGGGGTTTT  
13601 GCCTAGACTA GTGCTTAGTA GTAACTGTAT CACGGAAACCA CGGTCAAGGAC  
13651 TCTTGGCGTC CATCTGATCG TGGGAGACCC GTCAGCATGA GCTGGATCCC  
13701 CTCGGGGCCT GTCTTTCTT ACATAAAATGT TGCCCTTTGC CCTTACTTGG  
13751 TTTTTATTTT GTCCCGCGAC AATGGAAAAC TTAATTTTT TTTTTATTAA

## FIGURE 3F

13801 AAAGAAAAAT CTATTCTGGC CAGGTGCAGT GGCTCACGCC TGTAATCCCA  
13851 GCACTTGGG AGGCCAAGGC AGGCGGATCA CAAGGTCAAG AGATCGAGAC  
13901 CATCCTGGCT AACACAGTGA AACCCCGTCT CTACTAAAAA TACAAAAAAC  
13951 TTAGCCGGC GTGGTGGCGG GCGCCTGTAG TCCCAGCTAC TCGGGAGGCT  
14001 GAGGCAGGAG AATGGTGTGA ACCCAGAAGG CAGAGCTTGC AGTGAGCCGA  
14051 GATCACGCCA CTGCACTCCA GCCTGGCGA CAAAGTGAGA CTCTGTCTCA  
14101 AAAAAAAAGAA AAAGAAAAAT CTATTCTAAAG TGAAGCAGTT TTTCCAGTA  
14151 GGTGGCAGAA CTAATGCCA TTATGCCATT TATAATTTTA AGTGATTAAA  
14201 GAGGAGTAGT ATGTAGTATA TGCAAGGTCT AGCTCTAACAGA GCAAGTGCAGT  
14251 ATAAATAGTA GAAACTGACC TGATATTACA GTATGAGAAA CATGAAGGGG  
14301 TTCTGTTTG TGAGCTCTAA ATTTATCTTC CATGTATACT TCAAGGCTCT  
14351 TCTCCCCAGT AGATTTTAT TCATCTGAAC TATAATTAGG TGCCCTTTTCA  
14401 CCATTCTGAA AATAATTGGA TCAAATGCAT TTTAAAGTCC AGGGTCTGAA  
14451 AGGTGGAGGA ATCCTTCTC TTTACTGTTT CTAATTAAA CTCCCTTTCA  
14501 TTTACTAGAT TTCAGTCATG TCCAGAATTTC ATCTTTCTA AAAGCTTTAA  
14551 TCTAGATTAA GAAATCTAA ATCTTTATT TATTTTTTT TCGTTGAAGT  
14601 GCCCTGATTT TGTTGGTGGT AAAGACTCCA TTAGTATCCA CTTATACATT  
14651 TCCCTGACTT TGCCCTGTAC CAAACCTTAC AGTATTACA TTGTAATGTT  
14701 GCAATAATAA TAGCTAACAT ATTAATACAC TGAATATTG CTGTGTGCCT  
14751 AAGCTAAGGA TTTAATTCTC TTAAAATCCT GTGAGGTATT TTATTTTACA  
14801 GAAAAAGAAA CTGCTTAAAG AAAGTAACCT ATCCAGGTCA CACAAGTAAC  
14851 AATTGCAGAG CTGGAGTTTC AGATGAGGGC TGGCTTGCAGC TGCCGCTACA  
14901 GAAAAGAGTG CCCTAGAAAT CGGTCACTT GCATTTCCCG ATTTAGTTT  
14951 AGCCAAATGA AAAATTCTT TTGGATTAT GAGTATAATC AGACAGTATA  
15001 CCTGTGAAAT TAAAGTATTG GACTCTTGC TTGAAATAAG TAGGTTAAAA  
15051 AGATTTGGGT GGCGGGCGC AGTGGCTCAC GCCTGTAAATC CCAGCACTTT  
15101 GGGAGGCTGA GGCAAGTAGA TCATTTGAGG TCAGGAGTTC GAGACCAGCC  
15151 TGACCAATAT GGGGAAACCT CGTCTCTACT AAAAATACAA AAATTAGCCG  
15201 GGCCTGGTGG TGCTGCCTG TAATACCAGC TACTTGGAGG CTGAGGCAGG  
15251 AGAACATCTT GAAGCCAGGA GGCAGAGGTT ACAGTGAGCT GAGATCACGC  
15301 CACTGCACTC CAGCCTGGC AACAGAGGCC GACTCTGTCT AACAAACAAAA  
15351 AAGATTTGGG AAAACACTTT ATTAATGAAG AGTTCTGAC AAAGTATTG  
15401 TTTTGGGAG AATTCTTATA ATTGCTTTG AATATTAGGG TGCTCCCTT  
15451 TCTCTCATTG TAAATTCTTAC AGAGACTTAA GCACAGAGAA TTTTATTAC  
15501 ATGCCTGTTA ATTAATGTGT ATAATCAGAT TTTAACTATA TTTAGTGAAT  
15551 ATTAAGATTG AGGTACAAAT CAAGCCCTTT ATAATTAAAC ATACACATTG  
15601 AGAACATTAA TAAAATATTA AAACATTAAA CTGCTCTTCT CACCCACTCC  
15651 AAGTCAAATA GCATTTTTC AGTCAGGTGT CTGGGAGCTC GATGCAAGAT  
15701 AACAAAATCT GGTCTCTGCC TCAGGGAAACA TGAAATCTGT TTGGGGAAAGC  
15751 CAGAGCAAAA ATAAAGGTTT TAATAGCAAG CTCTCACTAA CTGCCCCCTGG  
15801 AAATCCACCC CACATCCTCC AGGAAGCCTT TCTCTACCC CAGTGCCTC  
15851 AGGAGCTTCT CCAAGGCAGG CCCTTCCCAG AGCGCAGTGT GCTCCCCAGC  
15901 TCACAGGAGA TGCTCCCTAC ACGCTGCAGG AAAGTCCAGT GCCTGCAGCA  
15951 CAGGCTTCAG CAGCAGACTC GGGTTCTAGT CTCAGTCTGC TGATTCTAG  
16001 TTGTGGAACC TGAGCAGGCG AAGTTACTAA ACCTCTCTGT GCGTCAGCCT  
16051 CCCAGGCTCG TTGCTTCAGG CCGCAGTTAG GCTGTGTGAA CAGGAGAGTG

## FIGURE 3G

16101 GGGATGGAA CTAGGTATCT TAAAGCGGGG CAGAGTTGG ATGAGCGGGC  
16151 CACCCCTCGT ATAGTTAGGA GGAAGATGAC GGGAGGCATG GAAGCTGGGA  
16201 TAGCCATCCT GAGTCAGTGC TAATTCTGAC ACTTCAGAAC ATCGAGTCAG  
16251 TCTGACCTGC GAGTGANCTT TCATTGACCA CTTAGAAACT ATTAGCACCT  
16301 TGGACAAACT ACTTCTTTC AGACCTGGTT GCTTCATGTC TGGATGGGA  
16351 AAACTGATAC TTAACTTGCA GATAGTGGTG AATCAAAAGT AGTATATGTG  
16401 AAGTACTCAC ACACGTGGG GCATTCAAGGC ATCGTCCCAT CCTACTTCTA  
16451 CCTTTACAT ATTGTAATAT GAAAGCTAAA CCATTTCTCG ATGTGAGTC  
16501 GTTTAATCG GCTACATAGT GAGTGGCATT CGATTTAAA AATGTCAACT  
16551 TGGGATCTGT CACCATGCTA CTTACCATTT GTATGTCACA CTGTTGAAT  
16601 GTCGGACCTG GTTGTGTTT CTCCAGATGG TATGTTACTT ACGGTCATGA  
16651 ACTGATTTGG AAGAACAGGG AGCCGCTAGT GAAAATCTGG CATGAAATAA  
16701 GGACTAATGG CCCCCAAAAA GGAGGTGGCT CTAAGTAAAA CTGGGATTGG  
16751 ACAGTAGTGG TGCACTCTGGT CTTGCCGCC TGAGAGCCCC AGGAGACATC  
16801 GGCTAGAGTG ACCATGGCTA TGCTCCCGTC TGGAAGATGC CAGCATCTGG  
16851 CCTCCCACTG TTTTCAGCTG TGTCCCCCAG TCCGTGCTT TTTAGAATGT  
16901 GAATGATGAT AAAGTTGTGA AATAAAGGTT TCTATCTAGT TTGTAAGCAG  
16951 ATGTGTGTGT TCTCTCTTTA AGGGGCCGAC ACGGCTCTGG CATTGGCTT  
17001 TGGTTGTTGC ATTGACAGGA CCTGGGGAGA GTGCACCTG AAAGGCCTGA  
17051 TCAGAACATG AAGGCGCTGG TTGCGCTGTCT TTGGACCCCTC CAGTGCCTCT  
17101 GCTTAGCCTT CACTCTTCCT TGCCCTCCCCC TCCCCCTGGT TGGCTGCACA  
17151 TAAAAGTCAA GAGTATCCCC TCTCCAGCAC AATCTGAAAT AACAGCTGCA  
17201 GTATTTCTC AATTTTCAGG AAAGGTAGTG TTTTCTGGCA GTGAGTGGCA  
17251 TATACAAAAA GCTATTTTCA GGTGTTGCTT TCTAGGTTCA ATTTGTAGAT  
17301 AAATTAAGAG GTAGAAAGAA GTGATTGAGG TAAATTCAAGA CTTGAAATCT  
17351 GAGCCGAATT TTATCTCTG TTTGAAAGTG TTCTAATTGA AGCGTCTCAC  
17401 TGAAAATAGC AGATAGTGGC TGTCGTCGTC ACAGCCCTCA CTGTTGTGGA  
17451 ATTCAATGTTA CCCTCGTGC TGAGAATGAC ATCTAGGAAA TGCAGTTGA  
17501 GAGTATGTTT CTCCTGAAGT CATTACAGG AGAATTTTA GTCTTTGAT  
17551 GGCTTAAAAA TGTATACCA AGTCTTGAG CTTGCTCCTG GGAGGATCGA  
17601 AGGCCCTGAT TTCAGCCTCC TGTGGCCGAT CGGACTCAGG TTGTGTGCCG  
17651 TGGGGGATGG GAATGGCGGC TTTGGAAAAG GAGTGGGAGT GGTGCCACC  
17701 TCACCAAGCA AGTGAGAACT GCATGGCAGC ACGCGCCCAAG CACATAGAAA  
17751 TTGTCAGTA TTTGGCAGTC CTTCATATCC TTCTCCATC AGGCTGGACT  
17801 TGTTCTACT ATGATTACA GTTATTCTTC CCAGGCACAG GATTCTGTT  
17851 TAAACTCGTA TCACTTCTAG GGGAGAGAGT TATCTTAGCC ATCATTTC  
17901 CAGGGAGGAA ACGGCACACG TGGTGTAGGG GCACTGCCCA AGGTACAAAT  
17951 GCTTGCTCT GACATCTGCT AACAACTGCA ACACAGATGA GGCAAGATGC  
18001 GTTTCCAGA GATGGGATAG GAGGCTGAGT TCATAGGGAC ATTCCCTCTA  
18051 GAGCCCAACA TTAATTACA TCGTGCTTTG GGCAGACAG GCAAAGAGGC  
18101 AATGAAGACA TCTCTGTGTC CCTGTTTGT GACTGGAAA AAGTTAGAAG  
18151 TCCCTGTAGC ATCTCCTGGT CCCTAAAACC CCTCAATGCT GGAGCCTCTG  
18201 TGCATGGCCT GGGGAGGCCA GAACTGGCT GTGGCCGGAG AAGCCTTGCT  
18251 GTCCACAGCT CCCTCCTGAT TGCCCACGAG GGTGCTTCAC TTTCTCCTCT  
18301 TGGCTTCTCT GGGGACCCGC GATCACTGCC TTCAAGGCCA TGCACCTCCCT  
18351 GGCCCGTGGG CCTCTGGGC TGTGCCGCCT CCACTGGCAT CTGAAGTGTG

## FIGURE 3H

18401 GGGTACCTAG GAACATGCCG TGGCTGCCGT CTCCCTCATT CCATACACTT  
 18451 CTTGAGTGGG TGCACTTGCT GAAGCCTCAG TTATCTGTGA GGATTCTGAG  
 18501 CTCCAGACCC ACAGAACATCTC TCTGTACTCT TAGTAAATGT GTCTACTGCA  
 18551 ACACACGCAT GGTTCCAGGC TCTGGGACCA CCCCCCCCAGGC CTGCACAGGC  
 18601 CCCTCAAATA GCACTCGGCT TAAGGAGTGA CACGAGCAAT CGGTGAAGTC  
 18651 TGAAACCCGG AGCCATTGGA GATCTCCCTC TCTCGCCTCT TATTTCTAGA  
 18701 ATTCAGCCCC TCAGCCTTCC CAGTGCCTGT GACTCCGTGG TGGTCCTCAC  
 18751 TTCTTAGTCC CTGGACTGTT GAGCCTGTT TCCAGCTGG TCTCCAAAGC  
 18801 AACCCGTGC TTCTCCATAT GCCTGCCAGA GTGCTAAAAA CACGTCTGTC  
 18851 ATTCCTTGT TGTACCTGT GAAAAACTTT TATTTATTTG AGACAGGGTC  
 18901 TCTCTCTCTC TCTCTCGTCC AGGCTGGAGT TCAGTGGTGC AATCTAGATG  
 18951 GTCACTACAC TCAGGGAGTT GGGGATGGCT CAGAGCTGTT AACAGAGAGG  
 19001 GGACTGCCA GGAGGACCTG CGTGAGGGGT GGGGGTGGGA TGACAAGGAA  
 19051 CCAGCTCTGG GAGTTGAAAG ACCTGGATTG AAGTCTCAAC CCAAGCCCTG  
 19101 GCCAGCTCTG GGACCCCCGA CAAGTCGGCC TCACTCTCTG CCCCTCAGTG  
 19151 GGCTCCTGTG TAGATGGGA TAATGATGGC TTTATATCCT GAGAATGTGG  
 19201 GGAGGGGATT AAGTGGCAA AATACCTGAG AGTGCCTGCAGTGCCTGG  
 19251 CTCAGCAAAT GCCCTTGTTC CCTCCCTCCC TCTCCCCAGA ACCCCTCCTC  
 19301 CCCTCTTCT TCTTTTTTT TTTTTTTTG TGACCCAGAG TCTTGCTATG  
 19351 TTGCCAGGC TGGAGTGCAG TGGCACAATC TCGGCTCACT GCAACCTCCA  
 19401 CCTCCTGGCT TCAGGCAATT CTTGTGCCTC AGCCTCTCGA GTAGCTGGGA  
 19451 TTACAGGCAG GCACCATCAC GCCCGGCTAA TTTTTTTTT TTTTTTTGT  
 19501 AGTAGAAATG GGATTTACC ATATTGGCAG GATGTTCTCG ATCTCCTGAC  
 19551 CTCAGGTGAT CCACTCGCCT TGGCCTCCC AAGTGCCTGGG ATTATAGGTG  
 19601 TCAGGCCACTG CGCCCAGCCC CCATTGTTA TCTCCTCTTC CATTCTTGT  
 19651 GGGGACTTTT AAAGGAAAAA TCAGGTTGGT GGGCTGGGG AGGGCATAGC  
 19701 TGAGACCACC TTGAGGGCAC CAAGCTCACT GACCAC (SEQ ID NO:3)

#### FEATURES:

Start: 2002  
 Exon: 2002-2098  
 Intron: 2099-5692  
 Exon: 5693-5763  
 Intron: 5764-12510  
 Exon: 12511-12612  
 Intron: 12613-12746  
 Exon: 12747-12844  
 Intron: 12845-16626  
 Exon: 16627-16735  
 Stop: 16736

#### SNPs:

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
--------------	-------	-------	--------	------------------	-------	-------

FIGURE 3I

237	T	C	Beyond ORF(5')
783	G	T	Beyond ORF(5')
1187	C	T	Beyond ORF(5')
1227	-	A T	Beyond ORF(5')
1450	T	C G	Beyond ORF(5')
3925	C	T	Intron
5539	G	C	Intron
7220	T	C	Intron
7396	G	A	Intron
9048	A	C	Intron
9952	T	C	Intron
10197	G	A T	Intron
10245	C	G	Intron
10427	C	T	Intron
10583	T	C	Intron
10651	A	G	Intron
11125	G	A	Intron
12025	A	C	Intron
12391	T	G	Intron
13001	A	G	Intron
13147	A	G	Intron
13587	A	G	Intron
13681	T	G	Intron
14336	A	G	Intron
14729	A	G	Intron
15124	C	T	Intron
15907	A	G	Intron
16341	-	G T	Intron
16786	G	C	Beyond ORF(3')
17159	G	A	Beyond ORF(3')
17976	-	T C	Beyond ORF(3')
18001	G	A	Beyond ORF(3')
18021	G	T	Beyond ORF(3')
18022	A	G	Beyond ORF(3')
18042	T	G	Beyond ORF(3')
18375	C	T	Beyond ORF(3')
19244	T	C	Beyond ORF(3')

Context:  
DNA  
Position

237 CGAGGTTCTTCATGTTGGTCAGGCTGGCTCGAACTCCGACCTCAGGTGATCCGTCG  
CCTCAGCCTCCAAAGTACTGCTGGGATTACAGACGTGAGCCACCGCACCCGGCCTTAT  
CTTTCATTTTTTCATGTATTTCTTTATTTAATCACTTATCCAGAAACATATCCT

CGTCTTGACAGTGCTGTGGTGCCTGTGGTTCCAGAAGCTGGTGTGCTGTGTC  
[T, C]  
GTGGTTGAGGAAGTTGCCATGGAACGTGACAGAGGAAGCAGAGTAGTCGGTGCCTT  
TCAGCCTAGTAGGCAGGATCAGGGACCCATCTTGCCTCTTGCCTGAACCACAATT  
GAATAAAACACCAAAAGCCCTGACTGATCATGATCATAGCAATCCGATTTATGATCATG  
GCCAGACCATTCTCAGGCGTCTTACCCAAGATATCAATCACTGGGTATGACAACCTA  
GACCTAAGGGTGCACCTGGTAGTAAAGATGATTAACCTCTCCAAAGGAATCTAAGGAATCCAG  
783 AAGGGTGCACCTGGTAGTAAAGATGATTAACCTCTCCAAAGGAATCTAAGGAATCCAG  
AGCAACACGAATCACTGCTCTTCCCTATAGGGTAAACCTCCAAGACTCCAGTCCCTGT  
GAGGAGGCCTGCCGCTGCCCTCCAGGGTCCAGGCTCCACATTGGGAGGTGTACA  
CAGTGCTCTCGCTTCAATTGCCTGTGTATGATCCCTTCCATCTTGCATAAATG  
CTGCTCTCACCATCTTAAAGAGTCTGGTAATTATTACCAAAGGTGGTATAAT  
[G, T]  
CTGTCACAGTCCCTGCTAGTGAGACATCTGATACAACGTGGAATCAGTTCAACAAAAT  
GCAGTAAAATTTATTTAATGTACTACGGAGAAGAAAAATGCTACAGTTATAAGATG  
CATCCTGATTCAGATATTAAAATGGAAAAATGCTTAAGATCTGTGAAAATGTAGCT  
TCCTTCCCACCTCTCAAGTGGAGAGCAAAACTGGACAGACTAGAAATGCCAGGGCT  
AGCTGAGAACCTTACAGAATGAGCAACTGCGGAAGGCCACAGGTAAACCCGAGATGTAGAT  
1187 CTACCAAGTTATAAGATGCATCCTGATTCAGATATTAAAATGGAAAAATGCTTAAGAT  
CTGTGAAAATGTAGCTTCCCTTCCCACCTCTCAAGTGGAGAGCAAAACTGGACAGAC  
TAGAAATGCCAGGGCTAGCTGAGAACCTTACAGAATGAGCAACTGCGGAAGGCCACAGGT  
AACACCGAGATGTAGATCAGCTGCCAGGGACAAGACAAGAATGTTCTAAAGTAAATC  
CTCTTACCAAGTATGTTATTGAAATCAGTCCATTGGCATCGAAGAAGGTGAAAGTGCTA  
[C, T]  
TTGCCTGTCCTACAGAGACTGGAGGAATGACAAATGTTAAATTATTTAAATTCAACA  
AGTAGAGGAATACCTGCTATGTGAAGGAGTTGTGCAATTCTAAATTAAATATATTTT  
TGAAGTTGTAGTTTCAATAATAATTCTTATCTAAATGTAACAAGTTAATTATATTA  
TCGAATAAACCTCAATTCTGACTAACAACATCAACACTACAGAAAAAGGAAAGTC  
ACTCAACTCCCACATGTAACAGACTTGAAGCAGTTGAGGTTCTAAATTATCC  
1227 TGGAAAAAAATGCTTAAGATCTGTGAAAATGTAACCTCTTCCACCTCTCAAGTGGG  
AGAGCAAAACTGGACAGACTAGAAATGCCAGGGCTAGCTGAGAACCTTACAGAATGAG  
CAACTGCGGAAGGCCACAGGTAAACCCGAGATGTAGATCAGCTGCCAGGGACAAGACAAG  
AATGTTCTAAAGTAATCCTTACAGTATGTTATTGAAATCAGTCCATTGGCAT  
CGAAGAAGGTGAAAGTGCTACTTGCCTGTCCTACAGAGACTGGAGGAATGACAATGT  
[-, A, T]  
TAAATTATTTAAATTCAACAAGTAGAGGAATACCTGCTATGTGAAGGAGTTGTGCAATT  
CTAAATTAAATATATTTGAAAGTTGTAGTTCAATAATAATTCTTATCTAAAT  
GTAACAAGTTAATTATATTCTGAAATAACCTCAATTCTGACTAACAACATCAACA  
CTTACAGAAAAAGGAAAGTCACTCAACTCCCACATGTAACAGACTTGAAGCAGTTG  
AGAGGTTCTAAATTATCCCTGAATTCTATCACATGACTATTCTCAGACATGTTG  
1450 TCAGTCCTTATTGGCATCGAAGAAGGTGAAAGTGCTACTTGCCTGTCCTACAGAGACT  
GGAGGAATGACAAATGTTAAATTATTTAATTCAACAAGTAGAGGAATACCTGCTATGT

## FIGURE 3K

GAAGGAGTTGTGGCAATTATAAAATTATATTGGAGTTGTAGTTTCAATAA  
 TAATTCTTATCTAAAATGTAACAAGTTAATTATATTATCGAATAAACCTCAATTTCGTA  
 GTACTAACACATCAACACTACAGAAAAAGGAAAGTCACTCAACTCCCACATGTAACAC  
 [T, C, G]  
 ACTTTAGAAGCAGTTGCAGAGGTTCTAAATTATCCCTGAATTCCCTACATGACTAT  
 TTTCTCAGACATGTTGACCTCACCTACACAGATGACTCACATATGTTCCATAAGCTG  
 GCAGTAAGTTAAGAACGATACCATGCCCTGAGGAAAAAGAAGTAATGTTAGCTCTCTA  
 CTCTGGCAAAGAACCTAATTCTGTATATTACTCTGTCTTGGTTGGCTATTATAGA  
 CAATAAATTATTGATCTGATTATAATTGAGAAAAGTAAGCTCTAAAGAAGTAAAATA

3925 GCCTCCGAGTAGCAGGAATTACAAACGTGCGCCACCACACCTGGTAATTTTATATT  
 TTAATAGAGATGGGTTGACTATGTTGCCAGGCTGGCTTGAACCTCCTGACTTAGTGA  
 TCCGCCTGCCCTGGCCTCCAAAGTGTGGGATTACAGGCGTGGCCACTGCGTCCGGCC  
 TAATTAAAAGTTAAAATGGATAATTTTATTGGCTGTGTGTTCATGATTACAGAC  
 TATGTTCTCTCTGTAGAGGTCCTTGTCTCCAATGTTGCTGTCAACATTCACTGA  
 [C, T]  
 TTCTCCTTATTACATGGCAGCAAATATGTATGTTGTGGAGCTCTCTCAGCATA  
 GTGAACATTCTGGGTCAGAGCAGTTCATGGCAGTGTACCTATCTGAGGTAAATGCTT  
 TAATCTGGGGCCTTGAGAGTATAAGCACTTAAGCTATCTGCAGAACGGACAAAGGGA  
 ATGATTACTGCCATATTCTACACGTAGTGAGTGCTCAGAACATATTGTTCTCACAGTG  
 TATGTAGAGAAGGGAGCCACAGATTGGTGGAGATGTTGCCTTCTGTCATTGCTGA

5539 ATGAGTCTCATGTTAGTGAGGAAATGGTAAC TGAGAAGTGGAGTGAATGACCGTG  
 TCGCTCAGCAGATCATGCAGCAGGTCAAGACTTCTACATCCCCTGAAAGTCGCTGAAATGA  
 TAGGCAGGAGAAGTATTCTACATGCCGTACCTCACAGTGTACCTGAGATTGAAACCGACACT  
 GTTATCTGTGTAGAAATCAGAAATGAAAACCTTCTACATGGCTGGATGTGGTGCAC  
 GCCTGTAATCCAGCTACTCAGGAGGCTGGGGACAAGAATAACTGAAACCGGTAGGCA  
 [G, C]  
 AGGTTGCAGTGAGCCAAATTGTACCACTGCACCTCAGCAGCCGGGCGAAAGAGTGAA  
 CTCTGTCTAAAAAAAAAAGAAAAGAAAAAAAAGTAAACCAATTGTTATACCTC  
 ACTTAAATTATTGTAATGTGACTTGTTTCAGGTGTATTCCAATTGTCAGTTACG  
 TGGTAAAGTTGCCACAGGAAGATGGACCATCACTGGTGCAGTAAGTATTCTATTG  
 TAAATTTTTTAATTAAATTAAATTACTTGAAATAAGTTAGACTTAGAAGAAT

7220 AGAAAAAAATTTTTAAGTGTCTTGAGTTAATGGCAGATTCTGGCACATGGAA  
 ATCTTATGTAATATTCTTACACATTCACTGTTGACTTATTAAATACTAATTCTT  
 AAATGCATTCAAATAGGGATTCTTAAAGGAACCTAAAGGTCAATTGAAA  
 AGAATTCTTATGAAAATAACCACTCCCTAAATTGTATGTTCCCAAATTGTTACCT  
 TAATTCTCTAGTGAGGCCTGTTCTGCTGTGACCACATGCTTCTAACGCTCCTT  
 [T, C]  
 TTTCCCTCGTGGATGTTATTCTTACAAATTGCTCTGATATAATTATAT  
 TTCGAATCATATTGTCACCTCATTCAACAGCTAACGACCTAAATATGAAGGCAGTGAA  
 GACCACTAGGATGAATCAGAGACTCAGAATTGCAATTAGCTGGGGAGAAAACATGCACA  
 CATCTAACACACTGAAAGGAATGAGGATTCTAGAGGACTTGGGGCTCAAGAGT  
 GAAGAGACCTTCTAATTAGCTGAAAGGACCTGCGAGGGCATTGATGTGCTTGGAC

FIGURE 3L

7396	GAAAAGAATTCTTATGTAAAATAACCATTCCCTAATTGTATGTTCCCCAAATTGTTA CACTTAATTTCCTAGTGAGGCCTGTGTTCTGCTCTGACCATGCTTCTTAAGCCT CCTTTTTCCCTCGTGAATGTTATTTCCTTATACAATTGCTCTGATATAATTAA TATATTTCGAATCATATTGTCTACCTCATTCAACAGCTAAGCACCTAATATGAAGGCA GTGAAGACCACTAGGATGAATCAGAGACTCAGAATTGAATTAGCTGGGGAGAAAACAT [G,A] CACACATCTAATACACACTGAAAGGAATGAGGATTCTCTAGAGGACTTTGGGGCTCTAA GAGTGAAGAGACCTTCTAATTAGCTGAAAGGACCTGCGAGGGCATTGATGTGCTCTT GGACAGCTGTTGCCTCATCTTATAGATAAGAAACTGAAGTGCCTTAATGAAGTATG GCAGTAAGGTATTGGAGTTAGAGTGGGGTGAATCCTGGCTGCTACTTACGTGTGAT TTCTAGGACATATTACTGAACCTCTGAATTTCAGTTCCCTTATAAAATGGGATAA
9048	GGCTCTTGTCACTGCAGGGCAGGGATGGGAGCTGAGGGCGTCAGGCTACCTAGTGTGCC TCTGCTAATGTCGCTGTGGCTAGGAGGCAAGGGTGCCTCTTCCGCTGACACCGCCTG TTAGGCGTATTGGGATGCCTCATTACAGTGTGGCAAGGGTGGGAGTCTAGGCTCTGCTCA GCCTTGTGGCACCCGTTCTAAATATTGTCTAAAAGGTCTCTTGTAGGCTAT CTTTTTGGTCCTGACTAGAGAGAACATGTTGAGGGATGATCGATATGAGGCCAAAG [A,C] AAGCCCAGGGAACCTACACACACATTGATTGAATCTCAGGCTCCTAGCTGGCCGC TTCTCTCTCTTCTTACAGCCTCTACATTGTTCATATGTAACACCCAGGGTC TTAGCTGTACTTAGCTTGTAAAGCAGAGGGAGCAGATTCACTAAATTATAACCAA ATAAAGTTAAAAACATAAGTATGATAGATTGAAGATTATAGATAACAGAAAAATGTT TGTGAGCCCAGGCGCAGTGGCTCACACTGTAATCCCAGCACTTGGGAGGCCGAGGTGG
9952	ATTGATGGAGAACAAAGACCTCACCTCTCCATGGACCCACACCTCTAGGCTGTT GGATCAGGGTCATGACTCACTGACTTAAACTGTGTATGAATGTCAGCGTTCTGAGA AGAGAAGGGTCATTTCTAAATTCTCTTGACTCGAAAAAGTGAAGAAAGTCTC TCTGCAAGGGAGTAAGCCAAATATTGTCAAAAAACAAGTGTGATTATTAGACATA TAAATATTAAATTATATAAAAGCCACATCGAGAAAATTCTAGAAGGATGATGAACTG [T,C] GTATGTAATAATTACAATAAGTTATAATCACAAAAAAACAGCGTCCATGGAATTGTAC AGATAACGACAATTTCATAACAGATGGAGAATAATCATCTATGGAATAGTAGTTAGA AGAACTTCATAGAATTTCATAACAGATGGAGAATAATCATCTATGGAATAGTAGTTAGA TCTTGTTGCCAGGCTGGAGTGCAAAGGTGCGATCTGGCTCGCTACAACCTCTGCCTCC CGGGTTCAAGCGATTCTCCTGCCTCACCTCTGAGTAGCTGGGATTACAGGCATGCACC
10197	ATTTAAATTATATAAAAGCCACATCGAGAAAATTCTAGAAGGATGATGGAACGTGTAT GTAATAATTACAATAAGTTATAATCACAAAAAAACAGCGTCCATGGAATTGTACAGAT AACGACAATTTCATAACAGATGGAGAATAATCATCTATGGAATAGTAGTTAGAAGAA CTTCATAGAATTTCATAACAGATGGAGAATAATCATCTATGGAATAGTAGTTAGAAGAA GTTGCCAGGCTGGAGTGCAAAGGTGCGATCTGGCTCGCTACAACCTCTGCCTCCGG [G,A,T] TCAAGCGATTCTCCTGCCTCACACCTCTGAGTAGCTGGGATTACAGGCATGCACCAACAT GCCCAGCTAATTGTATTAGCAGAGACTGGGTTCTCATGTTGGTCAGGCTGGTC TCGAACTCCAGACCTCAGGTGATCTGCCGCCCTCAGCTCCAAAGTCTGGGATTACAG GTGTAAGCGACTGTGCCTGGCAGAACATTAGAATTAAATGCTCTTATATCAACTA

FIGURE 3M

ATCAAATTATAATTGCTTCATTTGGGGAAACGTGTAATTGATTGTTGGGGTTT

10245	GGAACGTGTATGAAATAATTACAATAAGTTATAATCAGCAAAAAACAGCGTTCATGG AATTGTACAGATAACGACAATTGGGTTAACAGATGGAGAATAATCATCTATGGAATAGT AGTTAGAAGAACCTCATAGAATTGGGTTGGGAGTGCAGGTCGATCTCGGCTCGTACAACCT CTGCCTCCGGGTTCAAGCGATTCTCCTGCCTAACCTCTGAGTAGCTGGGATTACAGG [C, G] ATGCACCACCATGCCAGCTAATTGTATTAGCAGAGACTGGGTTCTTCATGTTG GTCAGGCTGGCTCGAACCTCCAGACCTCAGGTGATCTGCCGCCAGCCTCCAAAGTC CTGGGATTACAGGTGTAAGCGACTGTGCCTGGCAGAACCTCATAGAATTAAATGCTCTT TTATATCAACTAATCAAATTATTTGCTTCATTGGGAAACGTGTAATTGGATTGTT TTTGGGTTGGGTTGAGATAAAGTGTCACTCTGCGCCAGGCTGGAGTACAGTGGCTC [C, T] ATATCAACTAATCAAATTATTTGCTTCATTGGGAAACGTGTAATTGGATTGTT TTGGGTTGGGTTGAGATAAAGTGTCACTCTGCGCCAGGCTGGAGTACAGTGGCTCAA TCTGGCTCACCACAAACCTCAGCCTCCGAGTAGCTGGACTACAGGCGCCACACAC GTCTGGCTAATTGGTGTGTTAGTAGAGACGGGTTCACTATGTTGGCTAGGCTGGT CTTGAACCTGACCTCAGGTGATCCACCTGCCTGGCCCTCAGAGTGCTGGATTACA
10427	TTTCGTTCTTGTGCCCAGGCTGGAGTGCAGAACGGTGCATCTGGCTCGTACAACCTCT GCCTCCCGGGTTCAAGCGATTCTCCTGCCTAACCTCTGAGTAGCTGGGATTACAGGCA TGCACCACCATGCCAGCTAATTGTATTAGCAGAGACTGGGTTCTTCATGTTG TCAGGCTGGCTCGAACCTCCAGACCTCAGGTGATCTGCCGCCAGCCTCCAAAGTC TGGGATTACAGGTGTAAGCGACTGTGCCTGGCAGAACCTCATAGAATTAAATGCTCTT [C, T] ATATCAACTAATCAAATTATTTGCTTCATTGGGAAACGTGTAATTGGATTGTT TTGGGTTGGGTTGAGATAAAGTGTCACTCTGCGCCAGGCTGGAGTACAGTGGCTCAA TCTGGCTCACCACAAACCTCAGCCTCCGAGTAGCTGGACTACAGGCGCCACACAC GTCTGGCTAATTGGTGTGTTAGTAGAGACGGGTTCACTATGTTGGCTAGGCTGGT CTTGAACCTGACCTCAGGTGATCCACCTGCCTGGCCCTCAGAGTGCTGGATTACA
10583	AGAGACTGGGTTCTCATGTTGGTCAGGCTGGCTCGAACCTCAGACCTCAGGTGATCT GCCGCCTCAGCCTCCAAAGTCTGGATTACAGGTGTAAGCGACTGTGCCTGGCAGAA CTTCATAGAATTAAATGCTCTTATATCAACTAATCAAATTATTTGCTTCATTGG GGAAACGTGTAATTGGATTGGGTTGGGTTGAGATAAAGTGTCACTCTGCG CCAGGCTGGAGTACAGTGGCTCAATCTGGCTCACCACAAACCTCAGCCTCCGAGTAGC [T, C] GGGACTACAGGCGCCACCAACCGTCTGGCTAATTGGTGTGTTAGTAGAGACGGGG TTCACTATGTTGGCTAGGCTGGCTTGAACTCCTGACCTCAGGTGATCCACCTGCCTCG GCCCTCAGAGTGCTGGATTACAGGCGTGAGCCACCGTGCCCCGCTACAATTATAGTCT CTTGACAGAAGCCAGCTGGCTAAATTAGGTCTCTGGGCTCTTGGGAGT GTTCATGCTGCTTCCATCTGCAGTTACCTGACTTCAAGAATGCAACCCGAGCTTG
10651	CAGCCTCCAAAGTCCTGGATTACAGGTGTAAGCGACTGTGCCTGGCAGAACCTCATAG AATTAAATGCTCTTATATCAACTAATCAAATTATTTGCTTCATTGGGAAACG TGTAATTGGATTGGGTTGGGTTGGGTTGAGATAAAGTGTCACTCTGTCGCCAGGCT GGAGTACAGTGGCTCAATCTGGCTCACCACAAACCTCAGCCTCCGAGTAGCTGGGACTA CAGGCGCCACCAACCGTCTGGCTAATTGGTGTGTTAGTAGAGACGGGTTCACT [A, G] TGTGGCTAGGCTGGCTTGAACTCCTGACCTCAGGTGATCCACCTGCCTGGCCCTCA GAGTGCTGGATTACAGGCGTGAGCCACCGTGCCCCGGCTACAATTAGTCTCTTGACAC

## FIGURE 3N

		GAAGCCAGCTGGTCAAATTAGGTCTTCTGGGTCTCTTTGAGGAGTGTTCATGC TGTCTTCCATCTGCAGTTACCTGACTCTAAGAATGCAACCCGAGCTTGTTCCTG TTGAGGCCACTTGGCAGTTATATGAGGGACTGGGACATCTGAGATCTCTGGACTCATA
11125		TTCATGCTGCTTCCATCTGCAGTTACCTGACTCTAAGAATGCAACCCGAGCTTGT TTCCCTGTTGAGGCCACTTGGCAGTTATATGAGGGACTGGGACATCTGAGATCTCTGG ACTCATAATAATTTCTTAAAGTTAGTAATTCCCCAAATGTAAGATAATCTTGATT CTGAAGCAACCCGTACATAGAACATTAAGAAAACATTGATTAAGAGAGGTAGATGCT ATTTCCAGAAACAACCGTTTATATGAAAAGTAGGAAACCTTCTTTAATGATAGG [G, A] GCTTCTTCAAAAGTTATTTGCTCTTAGGTGTCTTTTTTTAAACATCTCATT CATAAATAATTAAAAACTTATGGAAAGTTGCAGGGAATAGTACAGAGGACTCCATAAA GTCTTTTTGTTGTTGTTGTTGTTGAGACAGAGTCTCGCTGTTACCCAGG CTGGAGTGCAGTGGACAATCTGGCTCACTGCAACCTCTGCCTCCGGGTTCAAGCAAT TCTCGGGCTTAGCATCTAAGTAGGTGGATTATAAGCATCCGCCACACGCCAGCTA
12025		AGCTTCCTAGTGGTCACTCCTCCTGCCCTCCTACCCCTGGCGACAACCTACCTACT TCTACTAAAGATAAATTAGTTGCAAATGGAACCATACAGCATATACTAGTATTGTTGT CCTGGCCTCATTACTCTGTATAATTACTTTGAGACTCATCCATGTTCTGTGTATCAGTT TATTCTTTATTATTTGAGACAGGGCTTACTCTGTTGCCAGGCAGGAGTGCAGTGG TGCAATCATAGCTCACTGTAACCTTGACCTCCTGGCTTAAGGGATCCTCATGCCTCACA [A, C] TGTGCTGGAATTACAGGCGTGAGCCACACACTGGAATGTTTGTCTTATGAAGAT GAATAAAGATTTCACATGAATTAAAGATGAAACATGCTTCATGCATGCAGGTTCTT TGGCGTATTGATGCCACTCCCTCTGGTTGGAGCTTGTCAAGAGAAGTGTGAGCAGTTC TTTCCTAGGCCATAGGTGAAAGATGCGCATGACACGCTTAGCACTGTCCTGCGGTTCAT GAGGCACATACATCTACTGCCCGTAGTAAAATTCACTCTTCCAAGCGATTACTGTG TGT
12391		AAGATTTCACATGAATTAAAGATGAAACATGCTTCATGCATGCAGGTTCTTGGC GTATTGATGCCACTCCCTCTGGTTGGAGCTTGTCAAGAGAAGTGTGAGCAGTTCTTCC TAGGCCATAGGTGAAAGATGCGCATGACACGCTTAGCACTGTCCTGCCGTTCATGAGGC ACATACATCTTACTGCCCGTAGTAAAATTCACTCTTCCAAGCGATTACTGTG GACATTAGTTCTTACCTATTATTGGGACATAAGTAACTGAAAGCTTGAAGCTTG [T, G] GCTCACCTAGAAATGTGCAGCATGTAACACTTCTAGAAAATGTGCTGCTCTTAGACCTT GTAGCCACTAAGCAGTTGCATATTGAGTTCCCATTCTCCCTGCTGTGTTACTTGCAGT CTGGTGCATCATGACAGTCTCGCAGCTGCTGCACTAAGATCCAGAAGGGAGGCTT CCATTATTCTTCCGATGTTACGTTACAGCAGGAAATGTAAGTATTATGAAGT GCAGTGCTGGGATAGTGGTGTGTTGAGTGGGTTCTGCCCTAAGTTAGA
13001		GCTGGAGCAATCACAGTTGCGCTTGTCTTGCTGCCTTCAGGCCCTGAAAGCCAT TATGCCATGGATAACAGCAGGAATGATCCTGGGATGGAATTTTGATCATGCGGCACA TCTTGGGGAGCTTTGGAATGTAAGTTGAGTGTAAATTGATTGCTAAACTGCTTCC TTGGGTCTGCGCTCTCTACCCAGCCTCACCCCTACCCCCATCCCCATGGCAGAGA CATTGAACATGCAACGGAAGCAGAAGCAGGTGGCTGGAGGGTGAGGAAACCTCAAC [A, G]

## FIGURE 30

	TGGCTTGCCTGGGTTACCCAGCATACTGGCTATTGTAGAGACAGTCTGTGCCCTTA CCCTACGCTTAACCTAACGGCCAACTGTTGGCCTGTTATTCCAGCCCCCTCTTAG AAGACTGCAGCCTGGCCCCAGTCTATGCTGACATCTCTTTCCCTCAGACTTCC TGCCCTCCTCTCCCTGCCTGGCCTCCACCCCTGCTACCCCTGACCTCTGCTCGCCAGTG CTATTTAGACATGCTGAGTTGGCGAGCATTGCTCTGTATGACTGGAGTAGAGGCCGGT
13147	AAGTTTGAGTGTAAATTGATTGCTAAACTGCTTCTGGTCATGCCCTCTACCCCA GCCTACCCCTACCCCCCATCCCCATGGCAGAGACATTGAACATGCAACGGAAGCAGAA GCAGGTGGCTTGGGAGGGTGAGGAAACCTAACATGGCTTGCTTGGGTTACCCAGCA TACCTGGCTATTGTAGAGACAGTCTGCTCTTACCCCTACGCTTAACCTAACGGT CAACTGTTGGCCTGTTATTCCAGCCCCCTCTAGAAGACTGCAGCCTGGCCCCAGTCT [A, G] TGCTGACATCTCTTCCCTCAGACTTCTGCCCTCTCCCTGCCCTGGCCTC CCACCCCTGCTACCCCTGACCTCTGCTCGCCAGTGCTATTAGACATGCTGAGTTGGCGGA GCCATTGCTCTGTATGACTGGAGTAGAGGCCGGTACTGCAAACCAATGTGGACCACTTA CTGAGTACCCGCTGTATGCAGGCACCAAGCTAGTCCCTATGTTAACTATTACTACTC CCATTTACTGATGGAAACTGAGGCTCAGACATCATCTCCCCAGGCCAACAGCTT
13587	GGAGTAGAGGCCGGTACTGCAAACCAATGTGGACCACTACTGAGTACCCGCTGTATGC AGGCACCAAGCTAGTCCCTATGTTAACTATTACTACTCCCTTACTGATGGAAA CTGAGGCTCAGACATCATCTCCCCAGGCCAACAGCTCTCAATAGCAGAGCAGAGCTG TAAACCCACCTCTATAAGCCCTTCCACCCCCACACCCATATGGAATTGGTGTAAA CTGCTTCTGGGTACAGCAAATGGCATTGTGGTTACAAGACCTCCACGTGTGCTTCA [A, G] ACAATGGGTTTGCCTAGACTAGTGTCTAGTAGTAACGTATCACGGAAACACGGTCAG GACTCTGGCGTCCATCTGATCGTGGAGACCCGTCAAGCATGAGCTGGATCCCTCGGG CCTGCTTTCTTACATAATGTTGCCTTGCCTTACTGGTTTATTGTTCCGC GACAATGGAAAACCTTAAATTAAAAAGAAAATCTATTCTGCCAGGTGC AGTGGCTACGCCTGTAATCCAGCACTTGGAGGCCAGGCAGCGATCACAGGTC
13681	ACTACTCCCATTTACTGATGGGAAACTGAGGCTCAGACATCATCTTCCCCAGGCCAAC AGCTCTCAATAGCAGAGCAGAGCTGAAACCCACCTCTATAAGCCCTTCCACCCCCAC CACACCATATGAAATTGGTTGCTAAACTGCTTGGGTACAGCAAATGGCATTGTGG TTACAAGACCTCACGTGTGCTTAAACAAATGGGTTTGCCTAGACTAGTGTCTAGTA GTAACTGTATCACGGAAACACGGTCAGGACTCTGGCGTCCATCTGATCGTGGAGACCC [T, G] TCAGCATGAGCTGGATCCCTGGGGCCTGCTTTCTTACATAAAATGTTGCCTTGG CTTACTTGGTTTATTGTTCCGCAGAATGGAAAACCTTAAATTAAAAATTAA AAGAAAATCTATTCTGCCAGGTGCAGTGGCTCACGCCTGTAATCCAGCACTTGGGA GGCCAAAGGCAGGCCAGTACAAGGTCAAGGAGATGAGACCATCCTGGCTAACACAGTGAA ACCCGTCTACTAAAAACCTAGCCGGCGTGGTGGCGGGCGCTGTAGT
14336	CTTGCACTGAGGCCAGATCACGCCACTGCACTCCAGCCTGGCGACAAAGTGAGACTCTG TCTCAAAAAAAAAAGAAAATCTATTCTAAGTGAAGCAGTTTCCCAGTAGGTGG CAGAACTAAATGCCATTATGCCATTATAATTAAAGTGTAAAGAGGAGTAGTATGTA GTATATGCAAGGTCTAGCTAACAGCAGTGCAGTATAAAAGTAGAAACTGACCTGATA

## FIGURE 3P

TTACAGTATGAGAAACATGAAGGGTTCTGTTTGTGAGCTCAAATTATCTTCATGT  
[A, G]  
TACTTCAAGGCTTCTCCCCAGTAGATTTTATTCTACTGAACATAATTAGGTGGCCT  
TTTCCATTCTGAAAATAATTGGATCAAATGCATTTAAAGTCCAGGGCTGAAAGGTGG  
AGGAATCCTTCTCTTACTGTTCTAATTAAACTCCTTCTATTACTAGATTCAGT  
CATGTCCAGAATTCTACATCTTCTAAAGCTTAATCTAGATTAGAAATCTAAATCTT  
TATTATTTTTCGTTGAAGTGCCCTGATTTGTTGGTGGTAAAGACTCCATTAGTA  
  
14729 ATTTAAAGTCCAGGGCTGAAAGGTGGAGGAATCCTTCTTTACTGTTCTAATTAA  
AACTCCTTTCATTACTAGATTCAGTCATGTCCAGAATTCTACCTTCTAAAGCTT  
AATCTAGATTTAGAAATCTAAATCTTTATTATTTTCTGTTGAAGTGCCCTGAT  
TTTGTGTTGGTAAAGACTCCATTAGTATCCACTTACATTTCCCTGACTTGCCTCTG  
ACCAAACCTTACAGTATTACATTGACTGTGCAATAATAAGCTAACATATTAAAC  
[A, G]  
CTGAATATTGCTGTGCCTAAGCTAAGGATTAATTCTCTAAATCCTGTGAGGTAT  
TTTATTTCAGAAAAAGAAACTGTTAAAGAAAGTAACCTATCCAGGTACACAAGTAA  
CAATTGAGAGCTGGAGTTAGATGAGGGCTGGCTTGCCTGCGCTACAGAAAAGAGT  
GCCCTAGAAATCGGTATCTGCATTCCGATTTAGTTAGCCAAATGAAAAATTCT  
TTGGATTATGAGTATAATCAGACAGTACACCTGTGAAATTAAAGTATTGACTCTTG  
  
15124 GTAACTTATCCAGGTACACAAGTAACAATTGAGAGCTGGAGTTCTAGATGAGGGCTGG  
CTTGCCTGCCCTACAGAAAAGAGTCCCTAGAAATCGGTATCTGCATTCCGATT  
TTAGTTAGCCAAATGAAAAATTCTTTGGATTATGAGTATAATCAGACAGTACACCT  
GTGAAATTAAAGTATTGACTCTTGCTGAAATAAGTAGGTTAAAAGATTGGGTGGC  
CGGGCGAGTGGCTACGGCTGTAATCCAGCACTTGGGAGGCTGAGGCAAGTAGATCA  
[C, T]  
TTGAGGTCAAGGAGTTGAGACCAGCCTGACCAATATGGGAAACCTCGTCTCTACTAAAA  
ATACAAAAATTAGCCGGCGTGGTGGTGCATGCCCTGAATACAGCTACTTGGAGGCTGA  
GGCAGGAGAAATCACTGAAGCCAGGAGGGCAGAGGTTACAGTGAGCTGAGATCACGCCACT  
GCACTCCAGCCTGGCAACAGAGCGCAGCTGTCTAACACAAAAAGATTGGGAAAA  
CACTTATTAAATGAAGAGTTCTGACAAAGTGATTTTTGGGAGAATTATAATTG  
  
15907 TTTTAAAATATTAACATTAACGCTCTCACCACTCCAAGTCAAATAGCATT  
TTTCAGTCAGGTGCTGGGAGGCTCGATGCAAGATAACAAATCTGGTCTCTGCCTCAGGG  
AACATGAAATCTGTTGGGAAGCCAGAGCAAAATAAGGTTTAATAGCAAGCTCTCA  
CTAACTGCCCTGAAATCCACCCCCACATCCTCCAGGAAGCTTCTACCCCCAGTCACAG  
CCTCAGGAGCTTCCAAGGCAGGCCCTCCAGAGCGCAGTGTGCTCCCCAGTCACAG  
[A, G]  
AGATGCTCCCTACACGCTGCAGGAAAGTCCAGTGCCTGCAGCACAGGCTTCAGCAGCAGA  
CTCGGGTTCTAGTCTCAGTCTGCTGATTCTAGTTGTGGAACCTGAGCAGGCGAAGTTAC  
TAAACCTCTGTGCGTCAGCCTCCAGGCTGTTGCTCAGGCCAGTTAGGCTGTG  
GAACAGGAGAGTGGGATGGGAAGTAGGTATCTAAAGCGGGGAGAGTTGGATGAGCG  
GGCCACCCCTCGTATAGTTAGGAGGAAGATGACGGGAGGCATGGAAGCTGGGATAGCCAT  
  
16341 GCGTCAGCCTCCCAGGCTGTTGCTCAGGCCAGTTAGGCTGTGTAACAGGAGAGTG  
GGGATGGGAAGTAGGTATCTAAAGCGGGGAGAGTTGGATGAGCAGGCCACCTCGT

## FIGURE 3Q

	ATAGTTAGGAGGAAGATGACGGGAGGCATGGAAGCTGGATAGCCATCCTGAGTCAGTGC TAATTCTGACACTTCAGAACATCGAGTCAGTCTGACCTGCGAGTGGCTTCAATTGACCA CTTAGAAACTATTAGCACCTTGGACAAACTACTTCTTCAGACCTGGTTGCTTCATGTC [-,G,T] GCGATGGGAAAATGATACTTAACCTGCAGATAGTGGTAATCAAAAGTAGTATATGTGA AGTACTCACACACTGCGGAGCATTCCAGGCATCGTCCCACCTACTTCTACCTTTACATA TTGTAATATGAAAGCTAACCACTTCAGTGTGAGTCAGTTAACCGCTACATAGTGA AGTGGCATTGATTTAAAAATGTCAACTTGGGATCTGTCACCATGCTACTTACCAATTG TATGTCACACTGTTGAATGTCGGACCTGGTTGTTCTCCAGATGGTATGTTACTTA
16786	TCTCGATGTGAGTCAGTTAACCGCTACATAGTGGCATTGATTTAAAAATGT CAACTTGGGATCTGTCACCATGCTACTTACCATTTGTATGTACACTGTTGAATGTCGG ACCTGGTTGTTCTCCAGATGGTATGTTACCTACGGTCATGAACGTGGGAAAGAA CAGGGAGCCGCTAGTGAAATCTGGCATGAAATAAGGACTAATGGCCCCAAAAAAGGAGG TGGCTCTAAGTAAAATGGGATTGGACAGTAGTGGTCATCTGGCCTGCCGCTGAGA [G,C] CCCCAGGAGACATCGGCTAGAGTGACCATGGCTATGCTCCGCTGGAAAGATGCCAGCAT CTGGCCTCCACTGTTTCAGCTGTGTCCCCAGTCGTGTTAGAATGTGAATGA TGATAAAGTTGTGAAATAAGGTTCTATCTAGTTGTAAGCAGATGTGTGTTCTCTC TTTAAGGGGCCGACACGGCTCTGGCATTTGCTTGGTTGTTGCATTGACAGGACCTGGG GAGAGTGCACCCCTGAAAGGCCTGATCAGAACATGAAGGCCTGGTTGCCTGCTTGGAC
17159	TGTTTCAGCTGTGCCCCAGTCGTGTTAGAATGTGAATGATGATAAAGTTGT GAAATAAAGGTTCTATCTAGTTGTAAGCAGATGTGTGTTCTCTTTAAGGGCCG ACACGGCTCTGGCATTTGCTTGGTTGTTGCATTGACAGGACCTGGGAGAGTGCACCC TGAAAGGCCCTGATCAGAACATGAAGGCCTGGTTGCCTGTTGGACCCCTCCAGTCCT CTGCTTAGCCTTCACTCTCCTGCCTCCCCCTCCCTGGTTGGCTGCACATAAAAGTC [G,A] AGAGTATCCCTCTCCAGCACAATCTGAAATAACAGCTGCAGTATTTCTCAATTTCA GAAAGGTAGTGTGTTCTGGCAGTGAGTGCATATAACAAAAGCTATTTCAAGGTTTGCT TTCTAGGTTCAATTGAGATAAAATTAAAGAGGTAGAAAGAAGTGATTGGTAAATTCA ACTTGAAATCTGAGCCGAATTATCTCTGTTGAAAGTGTTCTAATTGAAAGCGTCTCA CTGAAAATAGCAGATAGTGGCTGTCGTACAGCCCTACTGTTGGAATTGTT
17976	AAAAGGAGTGGAGTGGGCCACCTCACCAAGGCAAGTGAGAACTGCATGGCAGCACGCG CCCAGCACATAGAAATTGTCAGTATTTGGCAGTCCTTCATATCCTTCTCCATCAGGCT GGACTTGTCTACTATGATTACAGTTACGTTATTCTCCAGGCACAGGATTCTGTTCAAAC TCGTATCACTCTAGGGGAGAGAGTTATCTAGCCATCATTGCTGCCAGCGAGGAAACGGC ACACGTGGTGTAGGGCACTGCCAAGGTACAATGCTTGCTGTGACATCTGCTAACAA [-,T,C] TGCAACACAGATGAGGCAAGATGCGTTCCAGAGATGGGATAGGAGGCTGAGTTCAAG GGACATTCCCTCTAGAGCCAAACATTAATTACATCGTGTGTTGGCAGACCAGGAAAG AGGCAATGAAGACATCTCTGTCCTGCTTGTGACTGGAAAAAGTTAGAAGTCCCTG TAGCATCTCTGGTCCCTAAACCCCTCAATGCTGGAGCCTGTGCTGATGGCCTGGGAG GCCAGAACCTGGCTGTGGCCGGAGAAGCCCTGCTGTCACAGCTCCCTCCTGATTGCCA

FIGURE 3R

18001 TCACCAAGGCAAGTGAGAACTGCATGGCAGCACGCCAGCACATAGAAATTGTCCAGTA  
TTTGGCAGTCCTCATATCCTCTTCATCAGGCTGGACTTGTCTACTATGATTACA  
GTTATTCTTCCCAGGCACAGGATTCTGTTCTAAACTCGTATCACTTCTAGGGGAGAGAGT  
TATCTTAGCCATCATTTGCCAGCGAGGAAACGGCACACGTGGTAGGGGACTGCCA  
AGGTACAATGCTTGCTCTGACATCTGCTAACAACTGCAACACAGATGAGGCAAGATGC  
[G,A]  
TTTCCAGAGATGGGATAGGAGGCTGAGTTCATAGGGACATTCCCTAGAGGCCAACAT  
TAATTCACATCGTCTTGGCAGACCAGGAAAGAGGCAATGAAGACATCTGTGTCC  
CTGCTTGTGACTGGAAAAAGTTAGAAGTCCTGTAGCATCTCTGGTCCCTAAACCC  
CTCAATGCTGGAGCCTCTGTCATGGCTGGGAGGCCAGAACCTGGCTGGCCGGAGA  
AGCCTTGCTGTCCACAGCTCCCTCTGATTGCCACGAGGGTGCTCACTTCTCCCTT  
  
18021 GCATGGCAGCACGCCAGCACATAGAAATTGTCCAGTATTGGCAGTCCTCATATCC  
TTCTTCATCAGGCTGGACTTGTCTACTATGATTACAGTTATTCTCCAGGCACAG  
GATTCTGTTCTAAACTCGTATCACTTCTAGGGGAGAGAGTTATCTAGCCATCATTTGC  
CAGCGAGGAAACGGCACACGTGGTAGGGGACTGCCAAGGTACAATGCTTGCTCT  
GACATCTGCTAACAACTGCAACACAGATGAGGCAAGATGCGTTTCCAGAGATGGGATAG  
[G,T]  
AGGCTGAGTTCATAGGGACATTCCCTAGAGGCCAACATTAATTCACATCGTCTTGG  
GCAGACCAGGAAAGAGGCAATGAAGACATCTGTGTCCCTGCTTGTGACTGGAAAAA  
AGTTAGAAGTCCTGTAGCATCTCTGGTCCCTAAACCCCTCAATGCTGGAGCCTCTG  
GCATGGCCTGGGGAGGCCAGAACCTGGCTGTGGCCGGAGAACCTGCTGTCCACAGCTC  
CCTCTGATTGCCACGAGGGTGCTCACTTCTCCCTTGGCTCTGGGACCCCGCG  
  
18022 CATGGCAGCACGCCAGCACATAGAAATTGTCCAGTATTGGCAGTCCTCATATCCT  
TCTTCATCAGGCTGGACTTGTCTACTATGATTACAGTTATTCTCCAGGCACAGG  
ATTCTGTTCTAAACTCGTATCACTTCTAGGGGAGAGAGTTATCTTAGCCATCATTTGC  
AGCGAGGAAACGGCACACGTGGTAGGGGACTGCCAAGGTACAATGCTTGCTCTG  
ACATCTGCTAACAACTGCAACACAGATGAGGCAAGATGCGTTTCCAGAGATGGGATAGG  
[A,G]  
GGCTGAGTTCATAGGGACATTCCCTAGAGGCCAACATTAATTCACATCGTCTTGGG  
CAGACCAGGAAAGAGGCAATGAAGACATCTGTGTCCCTGCTTGTGACTGGAAAAA  
GTTAGAAGTCCTGTAGCATCTCTGGTCCCTAAACCCCTCAATGCTGGAGCCTCTG  
CATGGCCTGGGGAGGCCAGAACCTGGCTGTGGCCGGAGAACCTGCTGTCCACAGCTC  
CTCCTGATTGCCACGAGGGTGCTCACTTCTCCCTTGGCTCTGGGACCCCGCGA  
  
18042 ACATAGAAATTGTCCAGTATTGGCAGTCCTCATATCCTCTTCATCAGGCTGGACTT  
GTTCTACTATGATTACAGTTATTCTCCAGGCACAGGATTCTGTTCTAAACTCGTAT  
CACTTCTAGGGGAGAGAGTTATCTAGCCATCATTTGCCAGCGAGGAAACGGCACACGT  
GGTGTAGGGGACTGCCAAGGTACAATGCTTGCTCTGACATCTGCTAACAACTGCAA  
CACAGATGAGGCAAGATGCGTTTCCAGAGATGGGATAGGAGGCTGAGTTCATAGGGACA  
[T,G]  
TCCCTCTAGAGCCAACATTAATTCACATCGTCTTGGGAGACCAGGCAAAGAGGCAA  
TGAAGACATCTCTGTGTCCCTGCTTGTGACTGGAAAAAGTTAGAAGTCCTGTAGCAT  
CTCCTGGTCCCTAAACCCCTCAATGCTGGAGCCTCTGCTGACATGGCCTGGGAGGCCAGA  
ACCTGGCTGTGGCCGGAGAACCTGCTGTCCACAGCTCCCTCTGATTGCCACGAGGG

FIGURE 3S

TGCTTCACTTCTCCTCTGGCTCTCTGGGGACCCGCGATCACTGCCTCAAGGCCATG

18375 GCTTTGGGAGACCAAGGCAAAGAGGCAATGAAGACATCTCTGTGTCCGTCTTGACT  
GGGAAAAAGTTAGAAGTCCCTGTAGCATCTCCTGGCCCTAAACCCCTCAATGCTGGAG  
CCTCTGTGCATGGCCTGGGGAGGCCAGAACCTGGCTGTGGCCGGAGAACCTTGCTGTCC  
ACAGCTCCCTCCTGATTGCCACGAGGGTGCTCACCTTCTCCTGGCTCTGGGG  
ACCCCGCATCACTGCCTCAAGGCATGCACTCCCTGGCCGTGGCCTCTGGGCTGTG  
[C, T]  
CGCCTCCACTGGCATCTGAAGTGTGGGTACCTAGGAACATGCCGTGGCTGCCGTCTCCC  
TCATTCCATACACTTCTGAGTGGGTGACTTGCTGAAGCCTCAGTTATCTGTGAGGATT  
CTGAGCTCCAGACCCCACAGAACTCTCTGTACTCTTAGTAAATGTGTACTGCAACACA  
CGCATGGTCCAGGCTCTGGGACCACCCCCCGCCCTGCACAGGCCCTCAAATAGCACT  
CGGCTTAAGGAGTGACACGAGCAATCGGTGAAGTCTGAAACCCGGAGCCATTGAGATCT

19244 CTAGATGGTCACTACACTCAGGGAGTTGGGATGGCTCAGAGCTGTTAACAGAGAGGGGA  
CTGCCAGGAGGACCTGCGTGAGGGGTGGGGTGGGATGACAAGGAACCAAGCTCTGGAG  
TTGAAAGACCTGGATTCAAGTCTCAACCCAAGCCCTGGCCAGCTCTGGGACCCGGACAA  
GTCGGCCTCACTCTGCCCTCAGTGGCTCTGTAGATGGGATAATGATGGCTT  
ATATCCTGAGAATGTGGGAGGGATTAAAGTGGCCAAAATACCTGAGAGTGCCTGACTCAG  
[T, C]  
GCCTGGCTCAGCAAATGCCCTTGTCCCTCCTCCCTCCCCAGAACCCCTCTCCCT  
TCTTCTCTTTTTTTTTTTGACCCAGAGTCTGCTATGTTGCCAGGCTGGA  
GTGCAGTGGACAATCTGGCTCACTGCAACCTCCACCTCCTGGCTCAGGCAATTCTG  
TGCCTCAGCCTCTGAGTAGCTGGGATTACAGGCAGGCACCATCACGCCGGTAATTT  
TTTTTTTTTTGTAGTAGAAATGGGATTTCACCATATTGGCAGGATGTTCTCGATCT

Chromosome map:

Chromosome 3